Functional identification of an *Arabidopsis* pectin biosynthetic homogalacturonan galacturonosyltransferase

Jason D. Sterling*†, Melani A. Atmodjo*‡, Sarah E. Inwood*, V. S. Kumar Kolli*§, Heather F. Quigley*, Michael G. Hahn*¶, and Debra Mohnen*‖

*Complex Carbohydrate Research Center and Departments of ‡Biochemistry and Molecular Biology and ¶Plant Biology, University of Georgia, 315 Riverbend Road, Athens, GA 30602-4712

Communicated by Christopher R. Somerville, Carnegie Institution of Washington, Stanford, CA, January 20, 2006 (received for review December 12, 2005)

Galacturonosyltransferases (GalATs) are required for the synthesis of pectin, a family of complex polysaccharides present in the cell walls of all land plants. We report the identification of a pectin GalAT (GAUT1) using peptide sequences obtained from *Arabidopsis thaliana* proteins partially purified for homogalacturonan (HG) α-1,4-GalAT activity. Transient expression of *GAUT1* cDNA in the human embryonic kidney cell line HEK293 yielded uridine diphosphogalacturonic acid:GalAT activity. Polyclonal antibodies generated against GAUT1 immunoabsorbed HG α-1,4-GalAT activity from *Arabidopsis* solubilized membrane proteins. BLAST analysis of the *Arabidopsis* genome identified a family of 25 genes with high sequence similarity to *GAUT1* and homologous genes in other dicots, in rice, and in *Physcomitrella*. Sequence alignment and phylogenetic Bayesian analysis of the *Arabidopsis* GAUT1-related gene family separates them into four related clades of GAUT and GAUT-like genes that are distinct from the other *Arabidopsis* members of glycosyltransferase family 8. The identification of GAUT1 as a HG GalAT and of the GAUT1-related gene family provides the genetic and biochemical tools required to study the function of these genes in pectin synthesis.

biosynthesis | cell wall | glycosyltransferase | polygalacturonic acid | pectic polysaccharide

Pectins are structurally complex plant cell-wall polysaccharides that contain 1,4-linked α-D-galactopyranosyluronic acid residues. Galacturonic acid (GalA) is the most abundant glycosyl residue in the three types of pectin present in all plant primary walls: homogalacturonan (HG), rhamnogalacturonan I (RG-I), and rhamnogalacturonan II (RG-II) (1). Pectin accounts for ≈35% of dicot and nongraminaceous monocot primary walls and ≈10% of the primary wall of grasses (2). Pectins are also present in the walls of gymnosperms, pteridophytes, and bryophytes as well as *Chara*, a charophycean algae that is believed to be the closest extant relative of land plants (3).

Numerous studies show that pectins contribute to the physical and biochemical properties of the wall (1) and are required for normal plant growth and development (4). A complete understanding of pectin function requires knowledge of pectin biosynthetic enzymes and their corresponding genes. Although the enzyme activities of proteins encoded by some pectin biosynthetic genes, particularly those involved in the synthesis of sugar nucleotides, have been elucidated (5–9), the activities of putative pectin biosynthetic glycosyltransferases [e.g., QUA1 (10) and NpGUT1 (11)] have not been definitively demonstrated.

HG is a polymer of α-1,4-linked GalA that accounts for ≈65% of pectin. α-1,4-Galacturonosyltransferase (GalAT) activity has been identified in numerous plants (12) and shown to be membrane-bound in all species studied. Work in pea localized HG:GalAT activity to the luminal side of Golgi vesicles (13), the same location as pectin synthesis (14, 15). The most extensive study of HG:GalAT activity was done in tobacco (16, 17), and characteristics of the tobacco enzyme are comparable with the activity described in other plant species.

In vitro, GalAT transfers GalA from uridine diphospho-GalA (UDP-GalA) onto endogenous acceptors to produce pectic products with molecular masses of 100 kDa to >500 kDa (13, 16). Establishment of conditions to recover detergent-solubilized GalAT activity from membrane fractions (17) and *in vitro* studies using radiolabeled substrate (17) or fluorescently tagged (18, 19) acceptors established that, *in vitro*, GalAT preferentially transfers GalA onto the nonreducing end (20) of HG oligosaccharide acceptors [oligogalacturonides (OGAs)] of a degree of polymerization (DP) >9 (17, 18), although OGA acceptors as small as a trimer can be used (18, 19). Polymeric pectins, such as poly-GalA and pectin, are less favorable substrates (21). Membrane-permeabilized GalAT activity from pumpkin yielded a population of OGAs elongated by up to five galacturonosyl residues (19), whereas the solubilized petunia enzyme added up to 27 galacturonosyl residues onto the OGA acceptors (18). Clarification of the mode of action of GalAT(s) and the mechanism of HG synthesis requires access to purified or recombinantly expressed enzyme(s).

A protein-purification approach was taken to identify an *Arabidopsis thaliana* gene encoding HG:GalAT, because no gene encoding an enzymatically verified GalAT has been identified in any organism. GalAT was partially purified from *Arabidopsis* suspension-cultured cells and bioinformatics together with peptide sequence data were used to identify two putative GalATs. Functional characterization of the protein encoded by one of these genes (At3g61130), by using numerous biochemical methods, provides compelling evidence that this protein is a HG:GalAT (16). We therefore named this gene galacturonosyltransferase 1 (*GAUT1*).

Results and Discussion

Identification of *GAUT1*. Solubilized membrane proteins were isolated from log-phase suspension-cultured *A. thaliana* cells (22). GalAT activity was partially purified by sequential SP-Sepharose, Reactive yellow 3, and repetitive UDP-agarose chromatography (see lane 5 in Fig. 4, and Table 1, which are published as supporting information on the PNAS web site). The

Conflict of interest statement: No conflicts declared.

Abbreviations: DP, degree of polymerization; EPG, exopolygalacturonase; GalA, galacturonic acid; GalAT, galacturonosyltransferase; GATL, GAUT-like; HA, hemagglutinin; HG, homogalacturonan; HEK, human embryonic kidney; OGA, oligogalacturonide; RG-I, rhamnogalacturonan I; RG-II, rhamnogalacturonan II; UDP-GalA, uridine diphospho-GalA.

Data deposition: The sequence reported in this paper has been deposited in the GenBank database (accession no. DQ370437).

†Present address: Biolex Therapeutics, 158 Credle Street, Pittsboro, NC 27312.

§Present address: Windber Research Institute, 620 Seventh Street, Windber, PA 15963.

‖To whom correspondence should be addressed. E-mail: dmohnen@ccrc.uga.edu.

GalAT fraction was treated with trypsin, and resulting peptides were sequenced by liquid chromatography tandem MS. The peptide sequences obtained were screened against the *Arabidopsis* genome, revealing two related proteins (JS33 and JS36) that correspond to coding sequences At2g38650 and At3g61130, respectively (see Fig. 5*A* and Table 2, which are published as supporting information on the PNAS web site). Both sequences are predicted to encode glycosyltransferases with type II membrane topology (23), C-terminal catalytic domains with consensus sequences clustered in CAZy glycosyltransferase family 8 (24), and to have basic isoelectric points. These characteristics are consistent with the known *in vitro* biochemical properties of GalATs (17, 20). The predicted catalytic domains also contain DxD motifs, a conserved motif present in many well characterized glycosyltransferase families (25) and involved in the coordination of divalent cations (26). RT-PCR showed that transcripts from both genes are present in *Arabidopsis* flowers, roots, stems, and leaves (see Fig. 6, which is published as supporting information on the PNAS web site).

cDNA constructs containing N-terminally truncated versions of JS33 (JS33Δ1-43) and JS36 (JS36Δ1-41) downstream of a *Trypanosoma cruzi* mannosidase signal sequence, a polyhistidine tag, and two copies of a hemagglutinin (HA) epitope in the mammalian expression vector pEAK10 were generated (Fig. 5*B*) and used to transiently transfect human embryonic kidney (HEK)293 cells. Growth media from transiently transfected cells were treated with anti-HA antibodies bound to protein A-Sepharose. The protein immunoprecipitated from the JS36Δ1-41-transfected cells catalyzed the incorporation of [^{14}C]GalA from UDP-[^{14}C]GalA onto HG OGA acceptors (Fig. 1*A*). No GalAT activity was recovered from immunoprecipitates of the JS33Δ1-43 or empty vector-transfected cells. These results provided evidence that JS36 encoded a GalAT that catalyzes the transfer of GalA from UDP-GalA onto HG, thus, we named JS36 GAUT1.

Immunoabsorption of GalAT Activity from Partially Purified *Arabidopsis* SP-Fraction by Using GAUT1 Antiserum. To confirm that GAUT1 was involved in HG synthesis, GAUT1 antiserum was produced and used in an immunoprecipitation assay to deplete GalAT activity present in partially purified protein preparations from *Arabidopsis* suspension-cultured cells. The GAUT1 antiserum contained antibodies that recognized native (by immunoabsorption) and denatured (by Western analysis) recombinantly expressed truncated GAUT1 (see Fig. 7, which is published as supporting information on the PNAS web site) and an ≈60-kDa protein present in the partially purified *Arabidopsis* SP-Sepharose fraction. The GAUT1 antiserum immunodepleted GalAT activity from the *Arabidopsis* SP-Sepharose fraction (Fig. 1*B*, open diamonds). Furthermore, GalAT activity was immunoabsorbed by the conjugated GAUT1 antiserum in a dose-dependent manner (Fig. 1*B* black boxes), confirming that GAUT1 is a GalAT. Western analysis of the depleted fraction using increasing amounts of GAUT1 antiserum (see Fig. 8*A*, which is published as supporting information on the PNAS web site) demonstrated the progressive disappearance of the 60-kDa protein band from the SP-Sepharose fraction. Conversely, the 60-kDa band appeared in the immunoabsorbed fractions with increasing amounts of GAUT1 antiserum (Fig. 8*B*).

The ability of anti-GAUT1 immunoprecipitates to elongate OGAs was established by testing the sensitivity of the OGA products to cleavage by a HG-specific exopolygalacturonase (EPG) (Fig. 2 *A* and *B*). The bulk (98.9%) of the product synthesized by the anti-GAUT1 immunoprecipitates was cleaved by treatment with EPG (Fig. 2*A*, compare columns 7 and 8 with column 9), demonstrating that the synthesized products are HG. Furthermore, polyacrylamide-gel electrophoresis of the products generated by reacting anti-GAUT1 immunoprecipitate with

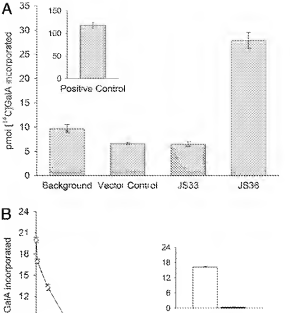


Fig. 1. GAUT1 has GalAT activity. (*A*) GalAT activity in equal amounts of media from HEK293 cells transiently transfected with the JS33Δ1-43 (JS33) or JS36Δ1-41 (JS36) cDNA constructs or empty vector control and immunoabsorbed with anti-HA antibodies conjugated to protein A-Sepharose. Solubilized *Arabidopsis* protein with GalAT activity (*Inset*) was used as a positive control. The average of time-0 reactions (background) is shown. Data are the average (±SE) of duplicate samples from 60-min reactions. Data were analyzed for significance compared with empty vector controls by using a two-sample, one-tailed Student's *t* test and an *a priori* α of 0.05 [95% confidence; $t_{S33} = 10.284$ ($P \sim 0.005$); $t_{S33} = 0.180$ ($P > 0.25$)]. Comparable results were obtained in three separate experiments. (*B*) Immunoabsorption of GalAT activity from *Arabidopsis* solubilized membrane proteins by using GAUT1 antiserum. Solubilized *Arabidopsis* membrane proteins partially purified by SP-Sepharose chromatography were incubated with increasing amounts of GAUT1 antiserum-coated Dynabeads, and proteins immunoabsorbed by anti-GAUT1 antibodies were magnetically separated from the SP-Sepharose fraction. GalAT activity was measured in both anti-GAUT1 immunoabsorbed (black boxes) and depleted (open diamonds) fractions and compared with similar fractions obtained by using preimmune serum (*Inset*). Open bar, preimmune serum immunodepleted fraction; filled bar, preimmune serum immunoabsorbed fraction.

UDP-GalA and trideca-GalA revealed the appearance of OGAs of increasing size (14-mer to 21-mer) in a time-dependent fashion (compare lanes 10–12 in Fig. 2*B*). Conversely, only a slight increase in the size of the OGAs was observed in the immunodepleted fraction (Fig. 2*B*, lanes 7 and 8). The sensitivity of the elongated OGAs made by anti-GAUT1 immunoprecipitates to treatment with EPG (Fig. 2*B*, lane 13) clearly demonstrates that the GAUT1 antiserum immunoabsorbs HG:GalAT activity.

GAUT1 and JS33 Are Part of a Multigene Family in *Arabidopsis*. BLAST analysis of GAUT1 amino acid coding region against the *Arabidopsis* genome identified JS33 and 13 additional coding sequences with 36–68% amino acid sequence identity and 56–84% sequence similarity to GAUT1 (genes *GAUT 1-15*; see Table 3,

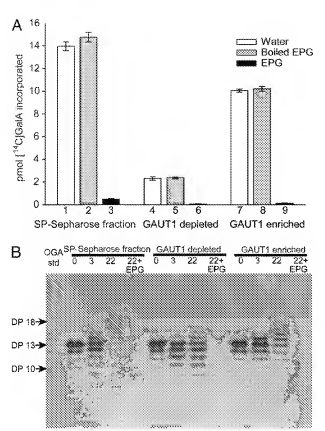


Fig. 2. Characterization of products made by anti-GAUT1 immunoabsorbed protein. (*A*) Sensitivity of synthesized product to cleavage by EPG. Products synthesized during 2-h reactions containing buffer, UDP-[^{14}C]GalA, OGA acceptors, and the *Arabidopsis* SP-Sepharose fraction, the SP-Sepharose fraction after immunodepletion of GAUT1 (GAUT1 depleted), or the anti-GAUT1 immunoabsorbed material (GAUT1 enriched). Each fraction was incubated overnight with water, boiled EPG, or native EPG. Radiolabeled products recovered by using the filter assay are shown. Note that small oligomers (e.g., monomer and dimer) do not bind to the filters. (*B*) Separation of nonradioactive GalAT reaction products by electrophoresis on a 30% polyacrylamide gel. Lanes 2–4 show the products recovered after incubation of SP-Sepharose purified solubilized *Arabidopsis* membrane proteins with UDP-GalA and OGAs enriched for a DP 13 for 0, 3, and 22 h at 30°C. Lanes 6–8 show the same series of incubation times with the SP-Sepharose fraction after immunodepletion with GAUT1 antiserum. Lanes 10–12 show similar series with the product synthesized by GAUT1-immunoabsorbed protein from the SP-Sepharose fraction. Lanes 5, 9, and 13 show product recovered after digestion of the respective 22-h reaction products with EPG. Lane 1 shows 0.1 μg of OGA standard of DP 7–23.

which is published as supporting information on the PNAS web site). An additional 10 sequences with reduced sequence identity to GAUT1 (23–29% identity and 42–53% similarity) were also identified and named the GalAT-like (*GATL*) genes (*GATL 1-10*). We suggest that all 25 *Arabidopsis* genes represent putative GalATs and have grouped them into a GAUT1-related gene superfamily (Table 3).

Sequence alignment of the 25 proteins by using CLUSTALX (27) revealed a high degree of alignment (see the first 25 entries in Fig. 9, which is published as supporting information on the PNAS web site) and the presence of several highly conserved domains (Fig. 3*A*) in addition to the DxD motif (25). The conserved GAUT1-related superfamily motif [H-[FWY]-[DNS]-G-x (2)-K-P-W-x (2)-[ILM]-[ADGS] is unique to the entire GAUT1-related superfamily. The *Arabidopsis* GAUT1-related superfamily members and orthologous proteins from rice, tobacco, and chickpea were the only sequences identified when this motif was used in a BLAST search of the SWISSPROT, TrEMBL, and BDP databases. Amino acid motifs were also identified that specifically discriminate the GAUT and GATL subclasses within the GAUT1-related superfamily. The GAUT family motif (H-x (2)-[ILV]-x-[ST]-D-N-[IV]-[IL]-A-[ASTV]-S-V-V-[AIV]-x-S-x-[AIV]-x (2)-[AS]-x (2)-[PS]-x (3)-V-[FL]-H-[ILV]-[ILV]-T-[DN]-x (2)-[NST]-x (2)-[AGP]-[IM]-x (3)-F) identifies the 15 *Arabidopsis GAUT* genes and their orthologues in other plants. The GATL motif W-M-x-[ILM]-Q-x (3, 4)-R-I-Y-[DEH]-L-G-S-L-P-P-F-L-L-[IV]-F-[AGS]-G-x-[IV]-x-[AP]-[IV]-[DENS]-H-[QR]-W-N-Q-H-G-L-G-G-D-N-[FILV]-x-[GS]-x-C-R identifies the 10 *Arabidopsis GATL* genes and their orthologues.

GAUT genes are all predicted to encode proteins with molecular masses between 61 and 78 kDa (Table 3). Most of the *GAUT* genes are likely to encode type II membrane proteins (23) that contain a putative transmembrane domain in their hypervariable N-terminal region (Fig. 3*A*). In agreement with this predicted topology, GAUT1 antiserum immunoabsorbs HG:GalAT activity, and such activity has been shown in pea to localize to the Golgi lumen (13). Furthermore, *Arabidopsis* GAUT1 protein has been localized to the Golgi (28). Three of the GAUT proteins (GAUT 3, 4, and 5) contain an N-terminal signal peptide rather than a transmembrane domain (Table 3). GAUT2 is the only member of the GAUT family that is predicted to contain no N-terminal transmembrane domain and no signal peptide. The GATL proteins have molecular masses between 39 and 44 kDa and are predicted to contain only a signal peptide at their N termini, suggesting that these proteins are not integral membrane proteins but are processed into the secretory pathway, a location consistent with a function in pectin synthesis.

The 25 proteins in the GAUT1-related superfamily belong to CAZy glycosyltransferase family 8 (24), which contains a total of 40 *Arabidopsis* proteins. We suggest that a splitting of CAZy family 8 be considered, based on functional and sequence-based analyses. Many of the current family-8 proteins are functionally annotated as being galactosyl- or glucosyltransferases, activities that are clearly distinct from the GalAT activity identified for GAUT1 and attributed to the other GAUT1-related proteins discussed here. Multiple sequence alignments of the *Arabidopsis* family-8 proteins revealed only a single region of sequence similarity across all 40 *Arabidopsis* proteins centered about the DxD motif. Even the DxD region common to the 40 *Arabidopsis* family-8 proteins contains differences that permit a distinction between the GAUT1-related family and the other family-8 proteins. For example, the GAUT1-related genes have a D-[DHS]-DxxxxxD motif, whereas the equivalent motif in the other family-8 proteins is D-[AG]-D that lacks the final D (Fig. 9). Furthermore, members of the GAUT1-related family have conserved amino acid motifs not found in any of the other *Arabidopsis* family-8 proteins (Fig. 3*A*). We believe that splitting family 8 as we suggest allows for evolutionarily more meaningful comparisons among this group of proteins.

Phylogenetic Analysis of the GAUT1-Related Superfamily. Phylogenetic analysis of the 25 GAUT1-related superfamily members show that they cluster into four distinct clades, with highly significant clade credibility values (Fig. 3*B*). The GAUT proteins cluster into three clades: GAUT-A (GAUT1–7), GAUT-B (GAUT8–12), and GAUT-C (GAUT12–15). All members of the GATL family cluster tightly into a distinct clade that is most closely related to GAUT15.

Our phylogenetic analysis of these proteins differs in significant respects from a published analysis (29), carried out by using CLUSTALW, particularly with respect to the phylogeny of the GAUT proteins. Our Bayesian analysis was restricted to the 25 GAUT1-related superfamily of proteins and included two proteins [GAUT11 (At1g18580) and GAUT5 (At2g30575)] not

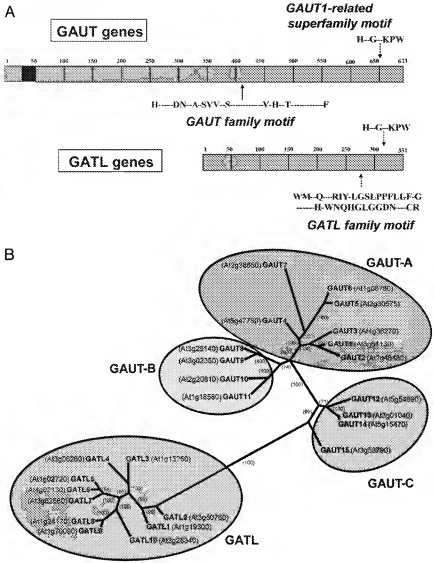


Fig. 3. Characterization of the *Arabidopsis* GAUT1-related gene superfamily. (*A*) Schematic representation of domain structure and conserved amino acid motifs in the *Arabidopsis* GAUT and GATL proteins. Positions of conserved residues found in the GAUT family are numbered relative to GAUT1. Residues 22–44 represent a predicted transmembrane region. The positions of conserved residues in the GATL family are numbered relative to GATL1. The conserved amino acid residues within the motifs diagnostic for the GAUT1-related superfamily and the GAUT and GATL subfamilies are shown. (*B*) Phylogenetic analysis of the GAUT1-related superfamily in *A. thaliana*. Alignment of the complete sequences of all 25 members of the GAUT1-related superfamily was carried out with CLUSTALX (27) using parameters suggested by Hall (44). Bayesian analysis employing the program MRBAYES (45, 46) was used to infer phylogenetic relationships among the members of the superfamily and to group the protein sequences into related clades. The analysis was carried out for 100,000 generations using a mixture of amino acid transition parameter models. The phylogram presented is the majority rule tree. Percentage clade credibility values for each branch are given in parentheses.

included previously. Our analysis also specifically excluded those family-8 proteins that show no substantial regions of sequence alignment with GAUT1. Bayesian analysis yields a different set of clades within the GAUT subfamily of proteins than observed previously. QUA1 and GAUT1 fall into clearly distinct clades in our analysis, whereas they had been grouped previously into the same clade. Functional characterization of additional members of the GAUT1-related superfamily should resolve which of the two models best describe the evolutionary relationships among these proteins.

A BLAST search of the National Center for Biotechnology Information (NCBI) (www.ncbi.nlm.nih.gov/entrez/query.fcgi) and TIGR (www.tigr.org) databases indicates the presence of several protein sequences from other plant species with significant identity to GAUT1. Phylogenetic analysis of available sequences from multiple species demonstrates that proteins with the most similarity to GAUT1 are present in soybean (TC207293), barrel medic (TC108344), tomato (TC154813), maize (TC272016), chickpea (CAB81547), and *Physcomitrella* (Contig3305). A search of the rice genome also reveals multiple proteins with high sequence similarity to members of the GAUT1-related superfamily. For example, the rice proteins Q652K2 and Q5Z7Z8 appear to be orthologues of GAUT1. Other gene sequences from rice (e.g., BAC06990) cluster into specific clades within the *Arabidopsis* phylogenetic tree. These results strongly suggest that GAUT1 and other members of the GAUT1-related superfamily are highly conserved in vascular and nonvascular plants.

Proposed Role in Pectin Synthesis. Pectin consists of HG, RG-I, RG-II, and, in some tissues and species, xylogalacturonan and apiogalacturonan (1). The precise nature of the linkages between these pectic polysaccharides in the wall remains controversial (30), although the available evidence supports a linkage via the backbones of the polymers (31). Clearly, the synthesis of pectin requires the coordinated action of numerous GalATs (at least one HG:GalAT, one RG-I:GalAT, and two to three RG-II:GalATs) (12). We propose that multiple members of the *GAUT1*-related gene family encode GalATs involved in different aspects of pectin synthesis. Consistent with this hypothesis, a survey of the *Arabidopsis* massively parallel signature sequencing (MPSS) database, the *Arabidopsis* microarray database GENEVESTIGATOR (www.genevestigator.ethz.ch) (32), and the whole-genome arrays of *Arabidopsis* (33) show that most members of the GalAT superfamily are expressed, at varying levels, in all of the major tissues of *Arabidopsis* (see Table 4,

which is published as supporting information on the PNAS web site).

Our recent analysis of walls from homozygous mutants of 12 members of the *GAUT1*-related gene family shows that mutants in 9 of the genes have significant reductions in the amount of GalA in their walls (K. Caffall and D.M., unpublished data), providing support for a function of other *GAUT1*-related genes as GalATs. Furthermore, plants carrying mutations in two members of the *Arabidopsis GAUT1*-related superfamily (*gaut8/qua1*) and (*gatl1/parvus/glz1*) have been described in refs. 10, 29, 34, and 35.

Qua1 mutant plants are dwarfed and have reduced cell adhesion and 25% reduced levels of GalA in their leaves. However, the pleiotropic effects of the mutation on glycosyltransferase activities make it difficult to specify the enzymatic function of the mutated gene (35). Our studies show that GAUT8/QUA1 is a member of the GAUT1-related superfamily and support the argument that GAUT8/QUA1 functions as a GalAT. Definitive proof will require expression of enzyme activity.

The *parvus/glz1* mutants are semisterile dwarf mutants. The neutral sugar compositions of *parvus/glz1* walls differ from wild-type walls in glycosyl residues found in RG-I. Because the levels of GalA in the *parvus/glz1* walls were not determined, it is not known whether *parvus/glz1* walls are altered in GalA content. Nevertheless, the phenotypes of plants carrying a mutation in the *GATL1/PARVUS/GLZ* and *GAUT8/QUA1* genes, together with the location of these proteins in different clades of the GAUT1-related superfamily, support the argument that GAUT8/QUA1 and GATL1/PARVUS/GLZ function as putative GalATs involved in pectin synthesis.

The identification of GAUT1 as a HG GalAT provides the molecular tools required to elucidate the biochemical mechanism(s) of HG and pectin synthesis. Biochemical and functional studies of GAUT1 and other members of the GAUT1-related gene family are expected to increase our understanding of the biological roles of pectin in plants and may lead to enhanced agricultural productivity and development of pectin-based pharmaceuticals and industrial polymers.

Materials and Methods

For a detailed version of this section, see *Supporting Materials and Methods*, which is published as supporting information on the PNAS web site.

Plant and Mammalian Cell Cultures. Cell-suspension cultures of *A. thaliana* (cv. Columbia) (22) were collected during exponential growth by filtration, washed extensively with water, and stored at −80°C until use. HEK293 cells (Edge Biosystems, Gaithersburg, MD) were grown in bicarbonate-buffered, Dulbecco's modified Eagle's medium (Sigma), pH 7.4, containing glucose (4.5 g/liter) and supplemented with 10% (vol/vol) FBS (Sigma), 0.6 g/liter L-glutamine, 100 μg/ml streptomycin sulfate, and 100 units/ml penicillin.

Preparation of SP-Sepharose-Purified *Arabidopsis* Membrane Proteins. Detergent-solubilized membrane proteins from suspension-cultured *Arabidopsis* cells (22) were loaded onto a SP-Sepharose column and bound proteins eluted by using a NaCl step gradient (see *Supporting Materials and Methods* for details). Proteins eluting with 300 and 400 mM NaCl were pooled, desalted, and stored at −80°C until use.

Partial Purification of GAUT1 and Identification by Liquid Chromatography Tandem MS. Desalted SP-Sepharose fraction was fractionated over Reactive yellow 3 and UDP-agarose columns, and the fraction most enriched for GalAT activity (see *Supporting Materials and Methods* and Table 1 for details) was treated with sequencing grade, modified trypsin (Promega). Resulting peptides were dialyzed overnight and analyzed by using a Q-TOF2 (Waters Micromass, Milford, MA) tandem mass spectrometer with a Waters CapLC delivery system. Survey MS spectra, acquired from 450–1,700 mass units, were used to identify peptide sequences. Sequences were probed against the *Arabidopsis* genome and proteins identified by using the Mascot search engine (www.matrixscience.com) and a peptide and tandem MS tolerances of ±2 and ±1, respectively.

Cloning of JS33 (GAUT7) and JS36 (GAUT1). A full-length cDNA encoding JS33 (GenBank accession no. AY091448) in vector pUNI51 was obtained from the *Arabidopsis* Biological Resource Center (ABRC, Columbus, Ohio). N-terminally truncated JS33 (JS33Δ1-43) lacking the putative transmembrane domain was generated by PCR using the sense primer FJS33Ntru (5′-ttattaccccgggaattccacaatggcttcactctcctggatttt-3′) and the antisense primer RpUNI51 (5′-ggcggctatgatctcgcggccgctagaattg-3′). The primers contained restriction sites for SmaI and NotI, respectively (underscored in the primer sequences).

A version of JS36 (GenBank accession no. NM_115977) lacking its putative N-terminal transmembrane domain was generated (JS36Δ1-41) from total flower RNA by using RT-PCR and the gene-specific sense primer FJS36Ntru (5′-ttattaccccgggaattccgaggagtgtatatcgatttcctcaaatg-3′) and antisense primer RJS36 (5′-gcggccgcattttattcatgaaggttgcaacgacg-3′). The primers contained restriction sites for SmaI and NotI, respectively (underscored in the primer sequences).

Expression of JS33 and JS36 in HEK293 Cells. Purified PCR products for truncated JS33 and JS36 were cloned into pCR2.1-TOPO (Invitrogen), and SmaI/NotI fragments were subcloned into the mammalian expression vector pEAK10 (Edge Biosystems) along with an N-terminal HindIII/SmaI fragment containing a *T. cruzi* mannosidase signal sequence, a polyhistidine tag, and two copies of the HA epitope (36). HEK293 cells (37) transiently transfected with 25–30 μg of DNA were incubated at 37°C for 48 h before harvesting.

Production of Polyclonal Antisera. Soluble, tetrameric multiple antigenic peptides corresponding to JS36 (GAUT1) amino acid sequences 448–472 (AMREYYFKADHPTSGSSNLKYRNPK) coupled to a trilysine core were synthesized at the Molecular Genetics Instrument Facility at the University of Georgia and used to produce polyclonal antibodies. Two rabbits (New Zealand White) were immunized (0.5 ml of 0.3 mgml^{-1} antigen), maintained and bled by the Polyclonal Antibody Facility at the University of Georgia.

Magnetic Bead Immunoabsorption of GalAT Activity. Solubilized SP-Sepharose-purified *Arabidopsis* membrane proteins containing GalAT activity were used for immunoprecipitation experiments (see *Supporting Material and Methods* for details). M-280 sheep anti-rabbit IgG-coupled Dynabeads (Dynal Biotech, Lake Success, NY) in PBS (6–7 × 10^8 beads per ml) were incubated for 2 h at 4°C with undiluted anti-GAUT1 antiserum at a ratio of 3:1 (vol/vol). Conjugate was washed and mixed with *Arabidopsis* SP-Sepharose fraction (0–160 μl of antiserum per 240 μl of SP-Sepharose fraction) by rotation for 2 h at 4°C. Washed immunoprecipitates and anti-GAUT1-depleted supernatants were analyzed by Western analysis and by using a GalAT filter assay (38). GalAT reactions containing 30 μl of enzyme, 50 mM Hepes, pH 7.3, 0.2 M sucrose, 0.05% (wt/vol) BSA, 25 mM KCl, 90 μg of OGAs (DP 7–23), 6.9 μM UDP-D-[^{14}C]GalpA (specific activity 196 mCimmol^{-1}) (1 Ci = 37 GBq), and 1.25 mM $MnCl_2$ in a total reaction volume of 60 μl were incubated for 2 h at 30°C. Reactions were terminated by the addition of 10 μl of 0.4 M NaOH.

Exopolygalacturonase Digestion of GalAT Products. Reaction product (total of 70 μl) was adjusted to pH 4.5 by the addition of 10 μl of 2 M acetic acid and 4.2 μl of 1 M sodium acetate buffer, pH 4.2. The mixture was incubated overnight at 30°C with 2 μl of water, a purified exopolygalacturonase (*Aspergillus tubingensis* EPG; EC 3.2.1.67, 0.54 mg/ml, 262 units/mg; 1 unit = 1 μmol of reducing sugar produced per minute), or EPG that had been incubated at >95°C for 1 hour. The digestion reaction was terminated by addition of 23 μl of 1 M NaOH. The final mixture was spotted onto cetylpyridinium chloride-coated filters and assayed by using the GalAT filter assay (38).

GalAT Filter Activity Assays. UDP-D-[^{14}C]GalpA was synthesized as described in refs. 39 and 40. The GalAT activity assay was a modification of the procedure of ref. 16 as described in ref. 38.

Nonradioactive GalAT PAGE Assay. The elongation of OGA acceptors by GalAT in the presence of UDP-GalA was assayed by separating reaction products on high-percentage acrylamide gels (see below). Reaction mixtures (15 μl) containing 0.33 μg/μl OGAs (DP 13), 3 mM UDP-GalA, 1.9 mM $MnCl_2$, 50 mM Hepes, pH 7.3, 200 mM sucrose, 25 mM KCl, 0.05% BSA, and 5 μl of enzyme was incubated at 30°C for 3 h (unless otherwise indicated). Reactions were terminated by the addition of 700 μl of chloroform/methanol (3:2) with vortexing and the mixture centrifuged for 5 min at 13,200 × g. The supernatant was removed and the pellet resuspended in 500 μl of 65% ethanol by using a combination of vortexing and sonication. The mixture was centrifuged for 5 min at 13,200 × g and the supernatant removed. The pellet was air-dried for 5 min, resuspended in 50 μl of sterile H_2O, and either frozen at −20°C or analyzed directly.

Samples were separated by PAGE and visualized by alcian blue/silver nitrate staining using a modification of the procedures of refs. 41 and 42 as described by ref. 43. Samples were mixed in a 5:1 ratio with sample buffer [0.63 M Tris·Cl, pH 6.8, 0.05% phenol red, 50% (vol/vol) glycerol], loaded onto a stacking gel [5% (wt/vol) acrylamide, 0.64 M Tris, pH 6.8] and separated over a 30% acrylamide resolving gel [0.38 M Tris, pH 8.8, 30% (wt/vol) acrylamide (37.5:1 acrylamide/bis-acrylamide, wt/wt)] at 17.5 mA for 60 min. The gel was stained for 20 min with 0.2% alcian blue in 40% ethanol and washed extensively with water (three times for 20 seconds and twice for 10 min). The gel was incubated in 0.2% silver nitrate containing 0.075% formaldehyde, rinsed three times for 20 seconds with water, and developed in 4% sodium carbonate containing 0.05% formaldehyde until bands appeared. The carbonate solution was removed, and staining was terminated by addition of 5% acetic acid.

This article is dedicated to the memory of Bruce P. Wasserman, a pioneer in plant cell wall biosynthesis research. We thank Maor Bar-Peled [University of Georgia (UGA)] for the *Arabidopsis* RNA; Stefan Eberhard (UGA) for the *Arabidopsis* suspension culture; Carl Bergmann (UGA) for the exopolygalacturonase; Kelley Moremen (UGA) for the pEAK vector and the HEK293 cell line; the *Arabidopsis* Biological Resource Center and SPP Consortium: Salk/Stanford/PGEC for cDNA clone U10739; and Malcolm O'Neill, Alan Darvill, and Maor Bar-Peled for critical reading of the manuscript. This work was supported by National Research Initiative, Cooperative State Research, Education, and Extension Service, U.S. Department of Agriculture Awards 2001353181-11111 and 2003-35318-15377 and, in part, by National Science Foundation Award 0090281 and Department of Energy Grant DE-FG05-93-ER20097.

1. Ridley, B. L., O'Neill, M. A. & Mohnen, D. (2001) *Phytochemistry* **57,** 929–967.
2. O'Neill, M., Albersheim, P. & Darvill, A. (1990) in *Methods in Plant Biochemistry*, ed. Dey, P. M. (Academic, London), Vol. 2, pp. 415–441.
3. O'Neill, M. A., Ishii, T., Albersheim, P. & Darvill, A. G. (2004) *Annu. Rev. Plant Biol.* **55,** 109–139.
4. Willats, W. G. T., McCartney, L., Mackie, W. & Knox, J. P. (2001) *Plant Mol. Biol.* **47,** 9–27.
5. Bonin, C. P., Potter, I., Vanzin, G. F. & Reiter, W.-D. (1997) *Proc. Natl. Acad. Sci. USA* **94,** 2085–2090.
6. Mølhøj, M., Verma, R. & Reiter, W.-D. (2003) *Plant J.* **35,** 693–703.
7. Burget, E. G., Verma, R., Mølhøj, M. & Reiter, W.-D. (2003) *Plant Cell* **15,** 523–531.
8. Gu, X. & Bar-Peled, M. (2004) *Plant Physiol.* **136,** 4256–4264.
9. Watt, G., Leoff, C., Harper, A. D. & Bar-Peled, M. (2004) *Plant Physiol.* **134,** 1337–1346.
10. Bouton, S., Leboeuf, E., Mouille, G., Leydecker, M. T., Talbotec, J., Granier, F., Lahaye, M., Höfte, H. & Truong, H.-N. (2002) *Plant Cell* **14,** 2577–2590.
11. Iwai, H., Masaoka, N., Ishii, T. & Satoh, S. (2002) *Proc. Natl. Acad. Sci. USA* **99,** 16319–16324.
12. Mohnen, D. (2002) in *Pectins and Their Manipulation*, eds. Seymour, G. B. & Knox, J. P. (Blackwell and CRC, Oxford), pp. 52–98.
13. Sterling, J., Quigley, H. F., Orellana, A. & Mohnen, D. (2001) *Plant Physiol.* **127,** 360–371.
14. Northcote, D. H. & Pickett-Heaps, J. D. (1966) *Biochem. J.* **98,** 159–167.
15. Moore, P. J., Swords, K. M. M., Lynch, M. A. & Staehelin, L. A. (1991) *J. Cell Biol.* **112,** 589–602.
16. Doong, R. L., Liljebjelke, K., Fralish, G., Kumar, A. & Mohnen, D. (1995) *Plant Physiol.* **109,** 141–152.
17. Doong, R. L. & Mohnen, D. (1998) *Plant J.* **13,** 363–374.
18. Akita, K., Ishimizu, T., Tsukamoto, T., Ando, T. & Hase, S. (2002) *Plant Physiol.* **130,** 374–379.
19. Ishii, T. (2002) *Plant Cell Physiol.* **43,** 1386–1389.
20. Scheller, H. V., Doong, R. L., Ridley, B. L. & Mohnen, D. (1999) *Planta* **207,** 512–517.
21. Takeuchi, Y. & Tsumuraya, Y. (2001) *Biosci. Biotech. Biochem.* **65,** 1519–1527.
22. Guillaumie, F., Sterling, J. D., Jensen, K. J., Thomas, O. R. T. & Mohnen, D. (2003) *Carbohydr. Res.* **338,** 1951–1960.
23. Reithmeier, R. A. F. & Deber, C. M. (1992) in *The Structure of Biological Membranes*, ed. Yeagle, P. (CRC, Boca Raton), pp. 337–393.
24. Henrissat, B., Coutinho, P. M. & Davies, G. J. (2001) *Plant Mol. Biol.* **47,** 55–72.
25. Wiggins, C. A. R. & Munro, S. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 7945–7950.
26. Breton, C. & Imberty, A. (1999) *Curr. Opin. Struct. Biol.* **9,** 563–571.
27. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **24,** 4876–4882.
28. Dunkley, T. P. J., Watson, R., Griffin, J. L., Dupree, P. & Lilley, K. S. (2004) *Mol. Cell. Proteomics* **3.11,** 1128–1134.
29. Lao, N. T., Long, D., Kiang, S., Coupland, G., Shoue, D. A., Carpita, N. C. & Kavanagh, T. A. (2003) *Plant Mol. Biol.* **53,** 687–701.
30. Vincken, J.-P., Schols, H. A., Oomen, R. J. F. J., McCann, M. C., Ulvskov, P., Voragen, A. G. J. & Visser, R. G. F. (2003) *Plant Physiol.* **132,** 1781–1789.
31. Nakamura, A., Furuta, H., Maeda, H., Takao, T. & Nagamatsu, Y. (2002) *Biosci. Biotechnol. Biochem.* **66,** 1301–1313.
32. Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. & Gruissem, W. (2004) *Plant Physiol.* **136,** 2621–2632.
33. Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M., *et al.* (2003) *Science* **302,** 842–846.
34. Shao, M., Zheng, H., Hu, Y., Liu, D., Jang, J.-C., Ma, H. & Huang, H. (2004) *Plant Cell Physiol.* **45,** 1453–1460.
35. Orfila, C., Sørensen, S. O., Harholt, J., Geshi, N., Crombie, H., Truong, H.-N., Reid, J. S. G., Knox, J. P. & Scheller, H. V. (2005) *Planta* **22,** 613–622.
36. Vandersall-Nairn, A. S., Merkle, R. K., O'Brien, K., Oeltmann, T. N. & Moremen, K. W. (1998) *Glycobiology* **8,** 1183–1194.
37. Jordan, M., Schallhorn, A. & Wurm, F. (1996) *Nucleic Acids Res.* **24,** 596–601.
38. Sterling, J. D., Lemons, J. A., Forkner, I. F. & Mohnen, D. (2005) *Anal. Biochem.* **343,** 231–236.
39. Liljebjelke, K., Adolphson, R., Baker, K., Doong, R. L. & Mohnen, D. (1995) *Anal. Biochem.* **225,** 296–304.
40. Basu, S. S., Dotson, G. D. & Raetz, C. R. H. (2000) *Anal. Biochem.* **280,** 173–177.
41. Corzo, J., Pérez-Galdona, R., León-Barrios, M. & Gutiérrez-Navarro, A. M. (1991) *Electrophoresis* **12,** 439–441.
42. Reuhs, B. L., Carlson, R. W. & Kim, J. S. (1993) *J. Bacteriol.* **175,** 3570–3580.
43. Djelinco, I. (2001) *Structural Studies of Pectin*, Doctoral Dissertation (Univ. of Georgia, Athens, GA).
44. Hall, B. G. (2004) *Phylogenetic Trees Made Easy: A How-To Manual* (Sinauer, Sunderland, MA), pp. 29–30.
45. Ronquist, F. & Huelsenbeck, J. P. (2003) *Bioinformatics* **19,** 1572–1574.
46. Huelsenbeck, J. P. & Ronquist, F. (2001) *Bioinformatics* **17,** 754–755.
47. Tavares, R., Aubourg, S., Lecharny, A. & Kreis, M. (2000) *Plant Mol. Biol.* **42,** 703–717.
48. Persson, S., Wei, H., Milne, J., Page, G. P. & Somerville, C. R. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 8633–8638.

Sterling *et al*. 10.1073/pnas.0600120103.

Supporting Information for

Sterling,J.D., Atmodjo,M.A., Inwood,S.E., Kolli,V.S.K., Quigley, H.F., Hahn,M.G., and Mohnen,D. (2006) Functional Identification of an *Arabidopsis* Pectin Biosynthetic Homogalacturonan Galacturonosyltransferase. *Proc. Natl. Acad. Sci. USA.* 103:5236-5241

Files in this Data Supplement:

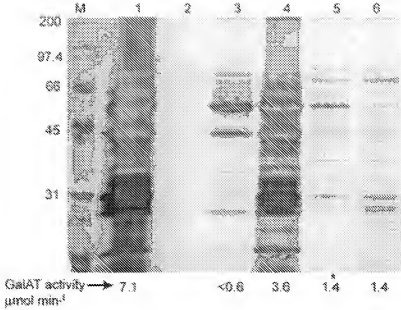


Fig. 4. SDS/PAGE of partially purified *Arabidopsis* solubilized membrane proteins. Fractions (20 μl) from the final UDP-agarose chromatography elution (see Table 1, step 5) were separated over a 10% SDS/PAGE gel and visualized by silver staining. Eluted fraction from the first UDP-agarose chromatography step is shown in lane 1; lanes 2–6 show eluted fractions from the second UDP-agarose chromatography step [flow-through (lane 2); 50 mM UDP (lane 3); 25 mM $MnCl_2$ (lane 4); 25 mM $MnCl_2$ and 0.5 M NaCl (lane 5); 25 mM EDTA (lane 6)]. Galacturonosyltransferase activity present in each fraction is shown below the figure. Molecular weight markers (M) in kDa are shown on the left. *, The fraction used for the identification of candidate galacturonosyltransferase proteins.

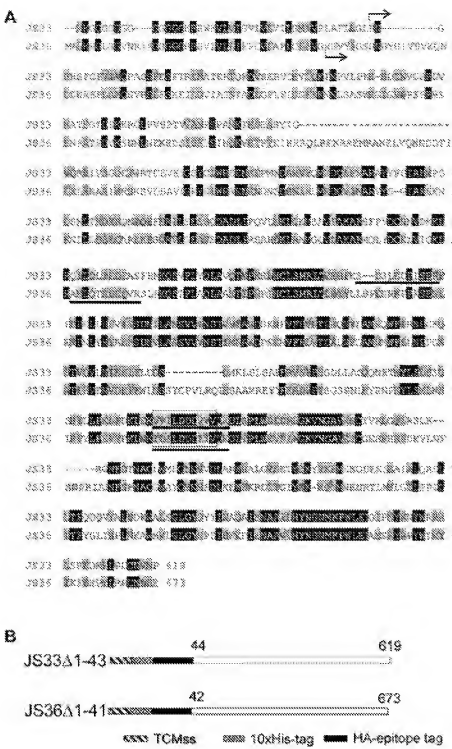


Fig. 5. Alignment of the amino acid sequences of JS33 (GAUT 7) and JS36 (GAUT 1). (*A*) The amino acid sequences of JS33 and JS36 were aligned by using CLUSTALX (1). The peptides identified by liquid chromatography tandem MS (LC-MS/MS) of tryptic peptides from the most purified GalAT-containing fraction from the UDP-agarose column are underlined. The arrows indicate the site of N-terminal truncation undertaken

to remove the N-terminal and predicted transmembrane regions. The conserved DxD motif (2) is boxed. (*B*) Schematic representation of the gene constructs used in the expression of the JS genes in HEK293 cells. The locations of the N-terminal *Trypanosoma cruzi* mannosidase signal sequence (TCMss), the polyhistidine tag (10× His-tag) and two copies of the hemagglutinin (HA)-epitope tag are indicated.

1. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **24,** 4876–4882.

2. Wiggins, C. A. R. & Munro, S. (1998) *Proc. Natl. Acad. Sci. USA* **95,** 7945–7950.

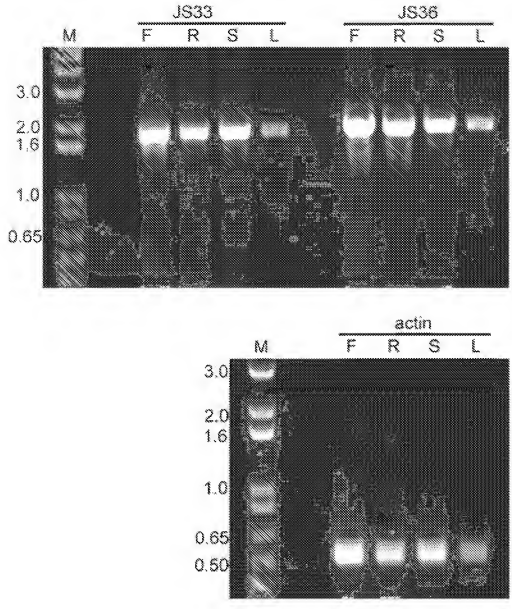


Fig. 6. RT-PCR expression analysis of JS33 and JS36 in *Arabidopsis*. Total RNA isolated from flowers (F), roots (R), stems (S), and rosette leaves (L) was used as a template for the RT-PCR of JS36 (GAUT1)- and JS33 (GAUT7)-specific transcripts. PCR products were separated on 1% agarose gels. The sizes of the amplified transcripts were 1,758 bp and 1,926 bp for JS33 and JS36, respectively. Primers directed against actin generated a 495-bp fragment that was used as a positive control. Positions of molecular weight markers (M) in kb are shown on the left.

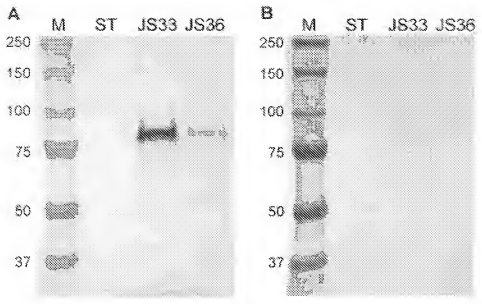


Fig. 7. Western blot analysis of media and cell lysates from transiently transfected HEK293 cells. JS33Δ1-43 (33) or JS36Δ1-41 (36) cDNA constructs were used to transfect 50% confluent (4-d-old) HEK293 cells. Media (*A*) or cell lysates (*B*) were obtained 48-h posttransfection, immunoprecipitated with 1 μg of anti-HA antibodies conjugated to 40 μl of protein A-Sepharose, separated on a 7.5% SDS/PAGE gel, and blotted onto PVDF membranes. The membranes were probed with a monoclonal anti-HA antibody because the recombinant JS33 and JS36 constructs contained an HA-epitope tag. Cells transfected with recombinant, histidine-tagged α-2,6-sialyltransferase (ST) lacking the HA epitope were used as a negative control. The positions of molecular weight markers (M) in kDa are shown on the left. The Western blot analyses show that the transiently transfected HEK293 cells expressed the recombinant enzymes in the media as expected. The estimated molecular weights of the recombinant JS33 (86 kDa) and JS36 (88 kDa) in both the media and cell lysates were much higher than those calculated from their amino acid sequences (73.2 kDa and 80.7 kDa for JS33 and JS36, respectively), suggesting that both proteins were posttranslationally modified in the HEK293 cells. JS36 was consistently expressed at lower levels in the transiently transfected and stably transformed HEK293 cells than JS33.

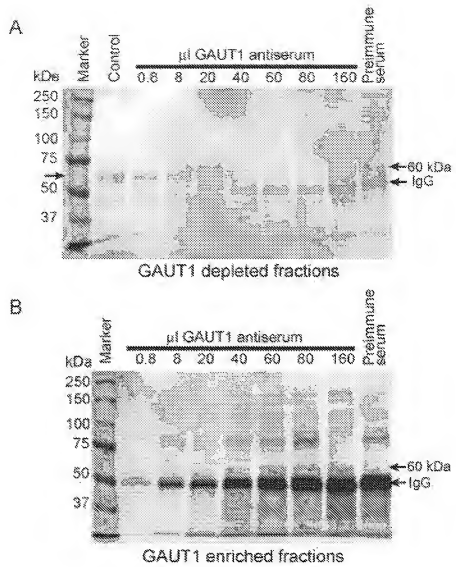


Fig. 8. Western blots of the GAUT1-depleted fractions (*A*) and GAUT1-enriched fractions (*B*) from the anti-GAUT1 immunoabsorption of GalAT activity from SP-Sepharose purified solubilized *Arabidopsis* membrane proteins described in Fig. 1*B* (see text). Equal proportions of depleted fractions and anti-GAUT1-conjugated magnetic beads were analyzed by immunoblotting with GAUT1 antiserum. Similar fractions obtained by using 160 µl of preimmune serum were used as controls. Arrows indicate the ≈60 kDa protein band detected by GAUT1 antiserum. IgG, IgG heavy chain detected by the secondary antibody used in Western blotting. Immunodepletion of GalAT activity from the SP-Sepharose fraction was accompanied by the disappearance of the 60-kDa protein band (*A*). Likewise, there was a dose-dependent appearance of both GalAT activity and the 60-kDa band in anti-GAUT1 immunoabsorbed fractions (*B*), suggesting that the 60-kDa band may represent a form of GAUT1. However, several other bands also appear to specifically react with the GAUT1 antiserum (*B*; compare immunoabsorbed and preimmune fractions) and the identity of these proteins remains to be determined. *GAUT1* encodes a predicted protein of 77.4 kDa; thus, the consistent

recognition and immunoprecipitation by GAUT1 antiserum of the 60-kDa protein in *Arabidopsis* fractions enriched for GalAT activity, and the correlation of the presence of the 60-kDa band with GalAT activity, could indicate that the 60-kDa protein is a proteolytically cleaved or otherwise processed version of GAUT1. The structure of the protein is under investigation.

Supporting Figure 9

Fig. 9. Sequence alignment of the 25 GAUT1-related superfamily proteins and other members of *Arabidopsis* glycosyltransferase family 8. The result of a multiple sequence alignment of 37 of 40 *Arabidopsis* family-8 proteins carried out by using CLUSTALX is shown. Three proteins (At2g35710, At4g16600, and At5g18480) were not included in the alignment because we were unable to find alignment parameters that would align the DxD motifs present in these three proteins with the corresponding motif in the other 37 family-8 proteins. The symbols below the sequence alignments are those defined in the CLUSTALX output. *, single conserved amino acid residue; :, fully conserved "strong" groups of amino acid residues; ·, fully conserved "weak" groups of amino acids. The first 25 proteins listed are members of the GAUT1-related superfamily. The first 15 are GAUT proteins and the next 10 are GATL proteins. GAUT1 (At3g61130) is in bold type. The remaining proteins listed in italics are other members of *Arabidopsis* glycosyltransferase family 8. None of the 15 members of family 8 that do not belong to the GAUT1-related gene family share significant sequence alignment with the GAUT1-related gene family proteins.

Table 1. Partial purification of enriched galacturonosyltransferase-containing *Arabidopsis* protein fraction used for tandem MS

Step	Sample	Total activity, μmol•min^{-1}	Yield, %
1	Solubilized membrane proteins*	57.2	100
2	SP-Sepharose†	37.5	65.6
3	Reactive yellow 3‡	17.5	30.6
4	UDP-agarose (1)‡	7.09	12.4
5	UDP-agarose (2)‡	1.35	2.4§

*Specific activity, 26 pmol•min^{-1}•mg^{-1}.

†Specific activity, 121 pmol•min^{-1}•mg^{-1}.

‡Because of high detergent and low protein levels, protein concentration not determined.

§Fraction used for tandem MS.

Table 2. List of proteins identified in the partially purified galacturonosyltransferase-containing fraction recovered after repetitive UDP-agarose chromatography

GenBank accession no.	Predicted molecular mass, kDa	Proposed function*	Mowse score†
BAB01930	65.9	β-fructofuranosidase	282
AAD05539	102	α-xylosidase precursor	236
AAL15216	42.0	putative methionyl-tRNA synthetase	168
BAB40450	76..6	long-chain acyl-CoA synthetase	112
AAD29817	61.7	synaptotagmin A	100
JS33/NP_565893	69.7	**putative glycosyltransferase**‡	110
BAB02117	48.5	unknown	81
BAB10278	35.7	unknown	77
JS36/NP_191672	77.4	**putative glycosyltransferase**	73
BAA20519	63.3	L-ascorbate oxidase	73
CAB78227	97.4	putative phospholipase D	71

*The proposed functions were previously inferred by electronic annotation.

†Mowse scores >60 are significant based on an *a priori* α of 0.05.

‡Proteins representing putative glycosyltransferases are in bold type.

Table 3. The *Arabidopsis GAUT1*-related gene superfamily

Gene*	GenBank accession no.†	Predicted molecular mass, kDa	Amino acid identity/similarity‡	SA versus SP§	Clade
GAUT1/JS36/LGT1	At3g61130	77.3	100/100	SA	GAUT-A
GAUT2/LGT2	At2g46480	62.1	65/78	none	GAUT-A
GAUT3	At4g38270	77.8	68/84	SP	GAUT-A
GAUT4/JS36L/LGT3	At5g47780	71.1	66/83	SP	GAUT-A
GAUT5/LGT5	At2g30575	69.9	45/67	SP	GAUT-A
GAUT6/LGT9	At1g06780	67.5	46/64	SA	GAUT-A
GAUT7/JS33/LGT7	At2g38650	69.7	36/59	SA	GAUT-A
GAUT8/QUA1	At3g25140	64.4	58/77	SA	GAUT-B
GAUT9	At3g02350	64.2	57/76	SA	GAUT-B
GAUT10/LGT4	At2g20810	61.8	50/72	SA	GAUT-B
GAUT11	At1g18580	61.9	51/71	SA	GAUT-B
GAUT12/LGT6/IRX8	At5g54690	60.9	40/61	SA	GAUT-C
GAUT13	At3g01040	61.1	43/62	SA	GAUT-C
GAUT14	At5g15470	65.3	43/62	SA	GAUT-C
GAUT15	At3g58790	60.6	37/56	SP/SA	GAUT-C
GATL1/PARVUS/GLZ1	At1g19300	39.0	29/49	SP	GATL
GATL2	At3g50760	32.5	27/52	SP	GATL

GATL3	At1g13250	39.9	23/43	SP	GATL
GATL4	At3g06260	40.3	29/51	SP	GATL
GATL5	At1g02720	41.2	25/44	SP	GATL
GATL6/LGT10	At4g02130	39.0	29/52	SP	GATL
GATL7	At3g62660	41.1	29/51	SP	GATL
GATL8/LGT9	At1g24170	44.0	23/42	SP	GATL
GATL9/LGT8	At1g70090	44.3	27/48	SP	GATL
GATL10	At3g28340	41.2	28/53	SP	GATL

*The name given to each member of the *GAUT1*-related gene family includes its designation within the LGT family [Tavares, R., Aubourg, S., Lecharny, A. & Kreis, M. (2000) *Plant Mol. Biol.* **42,** 703-717] or the names of any characterized *Arabidopsis* gene mutants (1–4). The numbering of the GAUT and GATL genes is based on the phylogenetic analysis of the family (see Fig. 3*B*).

†From The *Arabidopsis* Information Resource database or National Center for Biotechnology Information.

‡Sequence identity/similarity is compared to 397 amino acids of GAUT1 starting at amino acid position 277.

§The presence of either a putative transmembrane domain (SA; signal anchor) or a signal peptide (SP) sequence at the N-terminus of each gene was determined using SignalP version 3.0 (www.cbs.dtu.dk/services/SignalP) and confirmed using TMpred (www.ch.embnet.org/software/TMPRED_form.html).

1. Bouton, S., Leboeuf, E., Mouille, G., Leydecker, M. T., Talbotec, J., Granier, F., Lahaye, M., Höfte, H. & Truong, H.-N. (2002) *Plant Cell* **14,** 2577–2590.

2. Lao, N. T., Long, D., Kiang, S., Coupland, G., Shoue, D. A., Carpita, N. C. & Kavanagh, T. A. (2003) *Plant Mol. Biol.* **53,** 687–701.

3. Shao, M., Zheng, H., Hu, Y., Liu, D., Jang, J.-C., Ma, H. & Huang, H. (2004) *Plant Cell Physiol.* **45,** 1453–1460.

4. Persson, S., Wei, H., Milne, J., Page, G. P. & Somerville, C. R. (2005) *Proc. Natl. Acad. Sci. USA* **102,** 8633–8638.

Table 4. Expression analysis of the *GAUT1*-related superfamily

Gene	GenBank accession no.	Callus*	Inflorescence and early-stage floral buds*	Leaves*	Roots*	Siliques*	Whole-genome array†	GENEVESTIGA
GAUT1	At3g61130	+	+	+	+	+	+	+
GAUT2	At2g46480	–	–	–	–	–	–	+
GAUT3	At4g38270	+	+	+	+	+	+	+
GAUT4	At5g47780	+	+	+	+	+	+	+
GAUT5	At2g30575	+	+	+	+	+	–	NA§
GAUT6	At1g06780	–	+	+	+	–	+	+
GAUT7	At2g38650	+	+	+	+	+	+	+
GAUT8	At3g25140	+	+	+	+	+	+	+
GAUT9	At3g02350	+	+	+	+	+	+	+
GAUT10	At2g20810	+	+	+	+	+	+	+
GAUT11	At1g18580	+	+	+	+	+	+	+
GAUT12	At5g54690	–	+	+	+	+	+	+
GAUT13	At3g01040	+	+	+	+	+	+	+
GAUT14	At5g15470	+	+	+	+	+	+	+
GAUT15	At3g58790	+	–	–	–	+	+	+
GATL1	At1g19300	–	+	+	+	+	+	+
GATL2	At3g50760	+	+	+	+	+	+	+
GATL3	At1g13250	+	+	+	+	+	+	+
GATL4	At3g06260	–	–	–	–	+	+	+
GATL5	At3g62660	+	+	+	+	+	+	+

GATL6	At1g02720	–	+	+	+	–	+	+
GATL7	At4g02130	–	+	+	+	–	+	+
GATL8	At3g28340	+	+	+	+	+	+	+
GATL9	At1g24170	+	+	+	+	+	+	+
GATL10	At1g70090	+	+	+	+	+	+	+

*Presence or absence of MPSS signatures for specific galacturonosyltransferase superfamily genes according to the *Arabidopsis* MPSS database (http://mpss.dbi.udel.edu).

†Presence or absence of transcripts determined by whole genome array (1).

‡Presence or absence of transcripts in GENEVESTIGATOR database (2) determined by whole genome array.

§NA, gene not available in database.

1. Yamada, K., Lim, J., Dale, J. M., Chen, H., Shinn, P., Palm, C. J., Southwick, A. M., Wu, H. C., Kim, C., Nguyen, M., *et al.* (2003) *Science* **302,** 842–846.

2. Zimmermann, P., Hirsch-Hoffmann, M., Hennig, L. & Gruissem, W. (2004) *Plant Physiol.* **136,** 2621–2632.

Supporting Materials and Methods

Chemicals and Kits. Uridine diphosphate-α-D-[^{14}C]galactose (UDP-D-[^{14}C]Gal*p*, 11.0 GBq/mmol), PD-10 columns, and SP-, protein A- and protein G-Sepharose were purchased from Amersham Pharmacia Biosciences. A mixture of oligogalacturonides with a degree of polymerization (DP) of 7–23 was prepared according to refs. 1 and 3. UDP-agarose was purchased from Calbiochem (La Jolla, CA). Nickel-nitrilotriacetic acid agarose (Ni-NTA) resin, plasmid maxiprep, and gel extraction kits were purchased from Qiagen (Valencia, CA). EZNA plasmid miniprep kits were purchased from Omega Bio-tek (Doraville, GA). Restriction enzymes were purchased from New England Biolabs (Beverly, MA). Highly purified, monoclonal anti-hemagglutinin (anti-HA) antibodies were purchased from Covance (Berkeley, CA). Anti-HA conjugated to horseradish peroxidase was purchased from Roche. The 3,3',5,5'-tetramethylbenzidine (TMB) substrate kits were purchased from Vector Labs (Burlingame, CA). PVDF membranes were purchased from Schleicher & Schuell. Protein concentrations were determined by

the Bradford assay (Bio-Rad) using BSA as the standard. Reactive yellow 3 resin and all other chemicals were purchased from Sigma.

Plant Material. *Arabidopsis thaliana* (cv. Columbia) cell-suspension cultures derived from leaf calli were grown in the dark for 7–10 days in Gamborg's B-5 basal medium containing 58.4 mM sucrose, 9 μM 2,4-dichlorophenoxyacetic acid, and 0.23 μM kinetin as described in ref. 2.

Mammalian Cell Culture. Human embryonic kidney (HEK) 293 cells (Edge Biosystems; Gaithersburg, MD) were a gift of Kelley Moremen (University of Georgia, Athens). HEK293 cells were grown in 150-cm^2 culture dishes (Corning) in bicarbonate-buffered, Dulbecco's modified Eagle's medium (Sigma), pH 7.4, containing glucose (4.5 g/liter) and supplemented with 10% (vol/vol) FBS (Sigma), 0.6 g/liter L-glutamine, 100 μg/ml streptomycin sulfate, and 100 units/ml penicillin. Cells were maintained at 37°C in a humidified incubator (Thermo Electron, Newington, NH) containing 5% CO_2 and subcultured when the cells were 80% confluent.

Preparation of SP-Sepharose-Purified *Arabidopsis* Membrane Proteins. A detergent-solubilized membrane fraction from suspension-cultured *Arabidopsis* cells containing GalAT activity was prepared as described in ref. 2. The solubilized GalAT fraction (125 ml) was loaded twice onto a 50-ml column of SP-Sepharose cation exchange resin (5 × 2.5 cm) equilibrated with buffer A [50 mM Hepes, pH 7.3, 0.25 mM $MnCl_2$, 1% (vol/vol) Triton X-100, 2 mM EDTA, and 25% (vol/vol) glycerol]. The column was washed with 3 volumes of buffer B [50 mM Hepes, pH 7.3, 0.25 mM $MnCl_2$, 25% (vol/vol) glycerol, 2.5% (wt/vol) 3-[(3-cholamidopropyl)dimethylammonio]propanesulfonic acid CHAPS)]. Bound proteins were eluted by using a step gradient of NaCl in buffer B as follows: 50 ml of 100 mM, 50 ml of 200 mM, 35 ml of 300 mM, and 70 ml of 400 mM. Proteins eluting with 300 and 400 mM NaCl were pooled and desalted in buffer B over PD-10 columns according to the manufacturer's instructions. All operations were carried out at 4°C.

Reactive Yellow 3 Chromatography. The pooled and desalted fraction from the SP-Sepharose column was loaded onto a 15 ml column (2 × 5 cm) of Reactive yellow 3 (RY3) equilibrated with buffer B. Bound proteins were eluted by the sequential addition of one column volume of buffer B containing (*i*) 10 μM UDP, (*ii*) 10 μM UDP and 2 μM OGAs, (*iii*) 10 μM UDP, 2 μM OGAs and 100 mM NaCl, and (*iv*) 10 μM UDP, 2 μM OGAs, and 200 mM NaCl. The proteins that eluted with 100 and 200 mM NaCl were pooled and desalted in buffer B using PD-10 columns.

UDP-Agarose Chromatography. The pooled RY3 fractions were loaded onto a 7-ml (2.1 × 2 cm) UDP-agarose column equilibrated with buffer B. Bound proteins were sequentially eluted with one column volume each of 0.01, 0.05, 0.1, 0.5, 1, 5, 10, 15, 20 mM UDP and 1 M NaCl in buffer B. All of the eluted fractions were desalted and tested for GalAT activity. The fraction that eluted with 1 M NaCl and contained the greatest amount of GalAT activity was reapplied to the UDP-agarose column (Table 1) and bound proteins were eluted with one column volume of buffer B containing 0.01 mg/ml OGAs

and (*i*) 50 mM UDP, (*ii*) 25 mM $MnCl_2$, (*iii*) 25 mM $MnCl_2$ and 0.5 M NaCl, (*iv*) 25 mM EDTA, and (*v*) 25 mM EDTA and 0.5 M NaCl. OGAs were added to the elution buffers to inhibit the activity of a contaminating exopolygalacturonase-like activity that copurified with GalAT activity. The fraction that eluted with 25 mM $MnCl_2$ and 0.5 M NaCl was most enriched for GalAT activity and selected for further analysis by treatment with trypsin, followed by liquid chromotography tandem MS (LC-MS/MS).

Trypsin Digestion and LC-MS/MS Analysis of *Arabidopsis* Peptides. The most purified fraction from UDP-agarose chromatography was treated for 12 h at 37°C with 940 units of sequencing grade, modified trypsin (Promega). Resulting peptides were dialyzed overnight against water by using Spectro/Por CE 1,000 molecular-weight-cutoff dialysis membranes (Spectrum Laboratories, Rancho Dominguez, CA), lyophilized, and analyzed by LC MS/MS. The peptides were introduced into a Q-TOF2 (Waters Micromass, Milford, MA) tandem mass spectrometer using a Waters CapLC delivery system with gradient mobile phases of aqueous 0.1% (vol/vol) formic acid (A) and acetonitrile (B) and gradient conditions of 10% B to 70% B over 60 min at a flow rate of 1 μl/min. The Q-TOF2 mass spectrometer was operated in a data-dependent scan mode and survey MS spectra, acquired from 450-1,700 mass units, used to identify peptide sequences. The switch criteria for MS to MS/MS mode were ion count and charge state, with the Q-TOF set to ignore singly charged ions and to acquire MS/MS data for up to three coeluting peptides. Collision energy was varied automatically depending on the peptide mass and charge state. The sequences were probed against the *Arabidopsis* genome, and proteins were identified by using the Mascot search engine (www.matrixscience.com) and a peptide and MS/MS tolerance of ±2 and ±1, respectively.

RT-PCR. Total RNA from flowers, rosette leaves, or stems obtained from 3- or 6-week-old *A. thaliana* ecotype Columbia plants as well as total RNA from roots of 4-week-old *Arabidopsis* plants grown in liquid media were gifts of Maor Bar-Peled (University of Georgia) (4). First strand cDNA synthesis was conducted according to the manufacturer's instructions by using 1 mM oligo(dT) primer, 200 units of Superscript II reverse transcriptase (Invitrogen), and 8 μg of total RNA. Products from these reactions were treated with 2 units of RNase H (Invitrogen) and used as a template for PCR using gene specific primers. PCR reactions were conducted using high fidelity Platinum *Taq* polymerase (Invitrogen) according to manufacturer's instructions. Typical PCR reactions used 1 cycle at 95°C for 2 min, followed by 30 cycles at 95°C for 30 s, 52°C for 30s, and 70°C for 2.5 min and a final extension at 70°C for 5 min. The PCR products were analyzed on Tris-acetate EDTA (TAE) 1% (wt/vol) agarose gels, and DNA bands of the appropriate size were excised and purified by using a QIAquick gel extraction kit (Qiagen). Control PCR reactions included actin primers ACT119S and ACT284A (5) as a positive control [a gift from Richard Meagher (University of Georgia)].

Transcript analysis of JS33 (GAUT7) and JS36 (GAUT1) in *Arabidopsis* tissues was conducted by RT-PCR using the primers FJS33Ntru and RJS33 (5'-aagcattggccgtgtagctacttac-3') and FJS36Ntru and RJS36 for JS33 and JS36, respectively.

Expression of JS33 and JS36 in HEK293 Cells. Purified PCR products for truncated JS33 and JS36 were cloned into pCR2.1-TOPO (Invitrogen) and SmaI/NotI fragments were subcloned into the mammalian expression vector pEAK10 (Edge Biosystems) along with an N-terminal HindIII/SmaI fragment containing a *Trypanosoma cruzi* mannosidase (TCMss) signal sequence (MRLLTALFAYFIVALILAFSVSAKARR), a polyhistidine tag (HHHHHHHHHH), and two copies of the hemagglutinin (HA) epitope (YPYDVPDYA) (6). Gene constructs in the pEAK10 vector were confirmed by DNA sequencing.

Transient transfection and stable transformation of HEK293 cells (7) were carried out by using 25–30 μg of DNA. Transiently transfected cell lines were incubated at 37°C for 48 h before harvesting. Stably transformed cell lines were selected on media containing puromycin (1 μg/ml).

Stably transformed cell lines were confirmed by PCR amplification of genomic DNA and/or RT-PCR of total RNA using the gene specific primers FJS33Ntru and RJS33, and FJS36Ntru, RJS36, and 36IP (5'-tgattctgcagtccttggtaaatacag-3') for JS33 and JS36, respectively. Cells transformed with empty pEAK10 vector or expressing a recombinant, histidine-tagged, rat α-(2,6)-sialyltransferase (8), a gift of Kelley Moremen (University of Georgia), were used as negative controls. Control RT-PCR reactions conducted to test transcript levels in HEK293 cells included the use of primers (FP 5'-caatgacccctcattgacc-3', RP 5'-gtcttctgggtggcagtgat-3') specific to positions 123 and 584, respectively, of human glyceraldehyde 3'-phosphate dehydrogenase (accession no. BC023632) (gifts from Michael Pierce, University of Georgia).

Ni^{2+}-Affinity Column Chromatography. A recombinant rat α-(2,6)-sialyltransferase containing an N-terminal polyhistidine tag but no hemagglutinin epitope tag (gift of Kelley Moremen) was purified for use as a negative control for Western blot analysis of recombinant GAUT1 (JS36). Media from an HEK293 cell line expressing the recombinant α-(2,6)-sialyltransferase was desalted over a PD-10 column equilibrated with buffer C (50 mM KH_2PO_4, pH 7.5, 300 mM NaCl, and 40 mM imidazole). The desalted media (30 ml) was poured over a 200-μl column of Ni-NTA agarose equilibrated with buffer C and the column was washed with 10 volumes of buffer C. Proteins bound to the Ni-NTA resin were denatured in 0.2 volumes of SDS/PAGE sample buffer (9) and separated by SDS/PAGE.

Immunoprecipitation of Recombinant JS33 and JS36 Using Anti-HA Antibodies. Protein A-Sepharose, equilibrated in ice-cold PBS/137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 1.4 mM KH_2PO_4, pH 7.3), was incubated with 1 μg of purified monoclonal anti-HA antibody (Covance, Princeton, NJ) at a ratio of 30:1 (vol/vol) for 2 h unless otherwise specified. The resulting protein A-Sepharose/anti-HA conjugate was washed tree times with 10 volumes of cold PBS, brought to a 50% slurry with PBS (media) or cell-lysis buffer (cell lysates), and kept at 4°C until use.

Media (3.75 ml) from JS33- and JS36-expressing cell lines and vector alone cell lines, was harvested, brought to 10% (vol/vol) glycerol, and incubated with a 60-μl slurry of

protein A-Sepharose/anti-HA conjugate in the presence of 310 μl of PBS and 80 mg of BSA for 1 h. Proteins bound to the protein A-Sepharose/anti-HA conjugate were washed four times with 10 volumes of 0.1% (vol/vol) TX-100, 50 mM Tris•Cl, pH 7.4, and 300 mM NaCl followed by 2 washes with 10 volumes of PBS and used immediately for galacturonosyltransferase activity assays (10 μl) or SDS/PAGE (10 μl).

Cell lysates from HEK293 lines were prepared by incubating one 150-cm^2 dish of cells with 1 ml of cold cell-lysis buffer [50 mM Hepes, pH 7.3, 1% (vol/vol) TX-100, 150 mM NaCl, 2 mM EDTA, and 1 tablet of Complete Protease Inhibitor Mixture (EDTA-free; Roche)] for 30 min. Cell debris was pelleted by centrifugation, and the clear supernatant was incubated with 40 μl (for Western blots) or 160 μl (for galacturonosyltransferase activity assays) of protein A-Sepharose/anti-HA conjugate for 1 h. Immunoprecipitates were washed either five times with 10 volumes of buffer A (for activity assays) or three times with 10 volumes of buffer A, followed by 2 washes of 10 volumes of PBS (for Western blot). All manipulations were done at 4°C.

Production of Polyclonal Antisera. The criteria used to select the peptide sequence selected for MAP production and antibody production included uniqueness to GAUT1 (determined from sequence alignments using CLUSTALX, version 1.83) (10), low hydrophobicity (hydropathy plots) (11), high β-turn character (12) determined by using PROTSCALE (http://au.expasy.org/cgi-bin/protscale.pl) and significant loop character [determined using GENTHREADER (http://bioinf.cs.ucl.ac.uk/psipred)].

Magnetic Bead Immunoabsorption of GalAT Activity. A solubilized *Arabidopsis* membrane protein preparation (see above) containing GalAT activity (20 ml) was loaded onto a 5-ml column of SP-Sepharose equilibrated with buffer A [50 mM Hepes, pH 7.3, 0.25 mM $MnCl_2$, 2 mM EDTA, 25% (vol/vol) glycerol, 1% (vol/vol) Triton X-100]. The column was washed with 30 ml of buffer A, and bound proteins were eluted by using a step gradient of NaCl in buffer A (5 ml each of 100 mM, 200 mM, 300 mM, and 400 mM NaCl). The last 2.5 ml of the 300 mM NaCl fraction was pooled with the 400 mM NaCl fraction and desalted over a PD-10 column according to the manufacturer's instructions. This pooled, desalted fraction (referred to as the *Arabidopsis* SP-Sepharose fraction) was used for all immunoprecipitation experiments.

M-280 sheep anti-rabbit IgG-coupled Dynabeads (Dynal Biotech, Lake Success, NY) were washed with PBS (pH ≈7.3, 137 mM NaCl, 2.7 mM KCl, 4.3 mM Na_2HPO_4, and 1.4 mM KH_2PO_4), resuspended in PBS ($6–7 \times 10^8$ beads per ml), and incubated with undiluted anti-GAUT1 antiserum at a ratio of 3:1 (vol/vol) for 2 h at 4°C according to the manufacturer's instructions. The conjugate was washed three times in an excess volume of PBS and another three times in an excess volume of buffer A by using a magnetic particle concentrator (Dynal MPC-S; Dynal Biotech).

Immunoprecipitation was performed by mixing *Arabidopsis* SP-Sepharose fraction with increasing amounts of the Dynabead/anti-GAUT1 antisera conjugate (representing an equivalent of 0 to 160 μl of antiserum per 240 μl of SP-Sepharose fraction). The conjugate:SP-Sepharose fraction mixture was rotated for 2 h at 4°C, and

immunoprecipitates were washed either five times with an excess volume of buffer A (activity assays) or three times with buffer A, followed by two washes with PBS (Western blots).

GalAT assays of the SP-Sepharose fraction, the anti-GAUT1-depleted supernatants, and the anti-GAUT1 immunoabsorbed proteins (resuspended in buffer A) were done by using a filter assay (13). GalAT reactions containing 30 μl of enzyme, 50 mM Hepes, pH 7.3, 0.2 M sucrose, 0.05% (wt/vol) BSA, 25 mM KCl, 90 μg of OGAs (DP 7–23), 6.9 μM UDP-D-[^{14}C]Gal*p*A (specific activity 196 mCi•mmol^{-1}), and 0.125 mM $MnCl_2$ in a total reaction volume of 60 μl were incubated for 2 h at 30°C. Reactions were terminated by addition of 10 μl of 0.4 M NaOH.

SDS/PAGE and Immunoblotting. SDS-PAGE was carried out according to a method described in ref. 14 using 7.5% or 10% SDS/PAGE gels. Proteins were visualized by silver staining (15), Coomassie staining (Bio-Safe Coomassie G250 stain, Bio-Rad), or transferred to PVDF membranes for 2 h at 200 mA in 25 mM Tris, 192 mM glycine, and 20% (vol/vol) methanol. Membranes were blocked for 1 h in 3% BSA in Tris-buffered-saline [TBS; 100 mM Tris•Cl, pH 7.5, and 0.9% (wt/vol) NaCl] and incubated with a 1:1,000 dilution of anti-HA-peroxidase antibody in blocking buffer for 1 h. Membranes were washed four times in TTBS [0.1% (vol/vol) Tween 20 in TBS] and developed by using a 3,3',5,5'-tetramethylbenzidine (TMB) substrate kit (Vector Labs, Burlingame, CA) according to manufacturer's instructions.

Synthesis of UDP-D-[^{14}C]GalpA. UDP-D-[^{14}C]GalpA was synthesized based on methods described in refs. 16 and 17 with some modifications. Briefly, 50 μCi of UDP-d-[^{14}C]Gal*p* was purified over a CarboPac PA-1 column (Dionex; Sunnyvale, CA) by using a 0.05-1 M linear gradient of ammonium formate, pH 6.6, at a flow rate of 1 ml/min to ensure removal of contaminants that sometimes inhibit conversion of UDP-Gal to UDP-GalA. Purified UDP-D-[^{14}C]Gal*p* was lyophilized and resuspended in 0.5 ml of 25 mM NaH_2PO_4, pH 7.0, containing 250 units of galactose oxidase and 13,000 units of catalase. The reaction was incubated in the dark at room temperature for 24 h, after which another 250 units of galactose oxidase and 13,000 units of catalase were added. The reaction was incubated for another 24 h in the dark and terminated with 600 μl of 5:1 (vol/vol), chloroform/methanol. The resulting mixture containing UDP-D-[^{14}C]Gal*p*A was extracted and purified as described in ref. 16. The typical yield of UDP-D-[^{14}C]Gal*p*A from the starting UDP-D-[^{14}C]Gal*p* was between 60% and 70%.

Galacturonosyltransferase Filter Activity Assays. The galacturonosyltransferase activity assay used for determination of incorporation of radioactive [^{14}C]galacturonic acid from UDP-D-[^{14}C]GalA was a modification of the procedure of ref. 1 as described in ref. 13. Protein fractions (10 μl) were incubated in a total volume of 30 μl for 15 min at 30°C in reaction buffer [50 mM Hepes, pH 7.8, 0.2 M sucrose, 0.05% (wt/vol) BSA, and 25 mM KCl] containing 80 μg of oligogalacturonides (DP of 7–23), 1.1 μM UDP-D-[^{14}C]Gal*p*A (specific activity 298 mCi•mmol^{-1}), 0.25 mM $MnCl_2$, and 140 μM UDP-D-Gal*p*A. Assays of immunoprecipitated media from HEK293 cells were as above, except

that the reaction contained 2 μM UDP-D-[^{14}C]Gal*p*A and 100 μM UDP-Gal*p*A. Reactions were terminated by the addition of 5 μl of 0.4 M NaOH.

Terminated reaction mixtures were spotted onto cetylpyridinium chloride-coated Whatman 3-MM paper (4 cm^2) as described in ref. 13. The filters were allowed to dry for 5 min and were washed three times for 10 min each in 150 mM NaCl at a ratio of 1 filter per 100 ml of solution. The amount of radioactivity bound to each filter was estimated by scintillation counting. Statistical analysis of galacturonosyltransferase reaction data were performed by using a two-sample, one-tailed Student's *t* test (18) with an *a priori* α of 0.05 (95% confidence).

Analysis of Conserved Amino Acid Motifs. The sequence analysis program PRATT 2.1 (http://ca.expasy.org/tools/pratt) was used to identify conserved amino acid sequence patterns among the proposed 25-member *Arabidopsis* GAUT1-related superfamily and to distinguish the subgroups of that family: the 15 *GAUT* and 10 *GATL* genes. Conserved protein sequences were scanned against SWISSPROT, TrEMBL, and BDP by using SCANPROSITE, http://ca.expasy.org/tools/scanprosite. The presence of GAUT1 and other family members in the CAZy glycosyltransferase family 8 was determined by using the Carbohydrate Active Enzymes database at http://afmb.cnrs-mrs.fr/CAZY.

Sequence Alignment of GAUT Family Genes and Phylogenetic Analysis. *Arabidopsis* genes with high sequence identity to GAUT1 were aligned by using CLUSTALX (19) and parameters suggested by Hall (20) and phylogenetic analysis was done by Bayesian analysis employing MRBAYES (21). Phylogenetic analyses that included EST (GenBank, http://www.ncbi.nlm.nih.gov/dbEST) or tentative consensus (TC, http://www.tigr.org/tdb/tgi) sequences from soybean (*Glycine max*), barrel medic (*Medicago truncatula*), rice (*Oryza sativa*), tobacco (*Nicotiana tabacum*), and chickpea (*Cicer arietimum*) was conducted by using the program PAUP, version 3.1 (22).

1. Doong, R. L., Liljebjelke, K., Fralish, G., Kumar, A. & Mohnen, D. (1995) *Plant Physiol.* **109,** 141–152.

2. Guillaumie, F., Sterling, J. D., Jensen, K. J., Thomas, O. R. T. & Mohnen, D. (2003) *Carbohydr. Res.* **338,** 1951–1960.

3. Spiro, M. D., Kates, K. A., Koller, A. L., O'Neill, M. A., Albersheim, P. & Darvill, A. G. (1993) *Carbohydr. Res.* **247,** 9–20.

4. Harper, A. D. & Bar-Peled, M. (2002) *Plant Physiol.* **130,** 2188–2198.

5. McKinney, E. C., Ali, N., Traut, A., Feldmann, K. A., Belostotskyl, D. A., McDowell, J. M. & Meagher, R. B. (1995) *Plant J.* **8,** 613–622.

6. Vandersall-Nairn, A. S., Merkle, R. K., O'Brien, K., Oeltmann, T. N. & Moremen, K. W. (1998) *Glycobiology* **8,** 1183–1194.

7. Jordan, M., Schallhorn, A. & Wurm, F. (1996) *Nucleic Acids Res.* **24,** 596–601.

8. Wlasichuk, K. B., Kashem, M. A., Nikrad, P. V., Bird, P., Jiang, C. & Venot, A. P. (1993) *J. Biol. Chem.* **268,** 13971–13977.

9. Laemmli, U. K. (1970) *Nature* **227,** 680–685.

10. Thompson, J. D., Gibson, T. J., Plewniak, F., Jeanmougin, F. & Higgins, D. G. (1997) *Nucleic Acids Res.* **24,** 4876–4882.

11. Kyte, J. & Doolittle, R. (1982) *J. Mol. Biol.* **157,** 105–132.

12. Chou, P. & Fasman, G. (1978) *Adv. Enzymol. Relat. Areas Mol. Biol.* **47,** 45–148.

13. Sterling, J. D., Lemons, J. A., Forkner, I. F. & Mohnen, D. (2005) *Anal. Biochem.* **343,** 231–236.

14. Côté, F., Ham, K.-S., Hahn, M. G. & Bergmann, C. W. (1998) in *Plant–Microbe Interact.*, eds. Biswas, B. B. & Das, H. K. (Plenum, New York), pp. 385–432.

15. Heukeshoven, J. & Dernick, R. (1985) *Electrophoresis* **6,** 103–112.

16. Liljebjelke, K., Adolphson, R., Baker, K., Doong, R. L. & Mohnen, D. (1995) *Anal. Biochem.* **225,** 296–304.

17. Basu, S. S., Dotson, G. D. & Raetz, C. R. H. (2000) *Anal. Biochem.* **280,** 173–177.

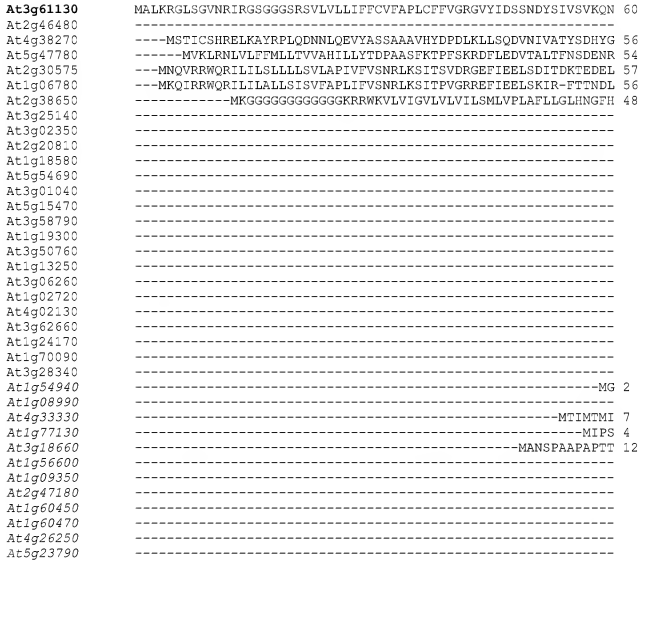
18. Zar, J. (1999) in *Biostatistical Analysis*, ed. Ryu, T. (Simon and Schuster, New York), pp. 122–158.

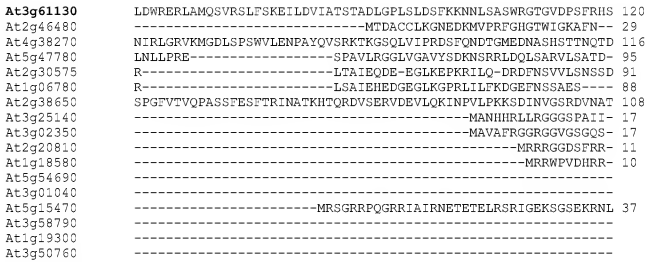
19. Jeanmougin, F., Thompson, J. D., Gouy, M., Higgins, D. G. & Gibson, T. J. (1998) *Trends Biochem. Sci.* **23,** 403–405.

20. Hall, B. G. (2004) *Phylogenetic Trees Made Easy: A How-To Manual* (Sinauer, Sunderland, MA).

21. Ronquist, F. & Huelsenbeck, J. P. (2003) *Bioinformatics* **19,** 1572-1574.

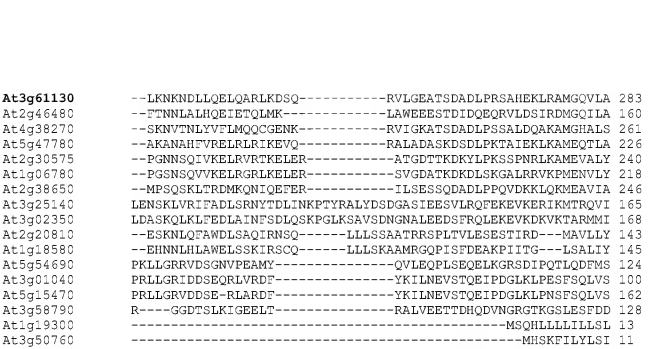
22. Swofford, D. (1998) PAUP. *Phylogenetic Analysis Using Parsimony and Other Methods* (Sinauer, Sunderland, MA).


At3g61130 MALKRGLSGVNRIRGSGGGSRSVLVLLIFFCVFAPLCFFVGRGVYIDSSNDYSIVSVKQN 60
At2g46480 ----------
At4g38270 ----MSTICSHRELKAYRPLQDNNLQEVYASSAAAVHYDPDLKLLSQDVNIVATYSDHYG 56
At5g47780 ------MVKLRNLVLFFMLLTVVAHILLYTDPAASFKTPFSKRDFLEDVTALTFNSDENR 54
At2g30575 ---MNQVRRWQRILILSLLLLSVLAPIVFVSNRLKSITSVDRGEFIEELSDITDKTEDEL 57
At1g06780 ---MKQIRRWQRILILALLSISVFAPLIFVSNRLKSITPVGRREFIEELSKIR-FTTNDL 56
At2g38650 -----------MKGGGGGGGGGGGGKRRWKVLVIGVLVLVILSMLVPLAFLLGLHNGFH 48
At3g25140 ----------
At3g02350 ----------
At2g20810 ----------
At1g18580 ----------
At5g54690 ----------
At3g01040 ----------
At5g15470 ----------
At3g58790 ----------
At1g19300 ----------
At3g50760 ----------
At1g13250 ----------
At3g06260 ----------
At1g02720 ----------
At4g02130 ----------
At3g62660 ----------
At1g24170 ----------
At1g70090 ----------
At3g28340 ----------
At1g54940 ----------MG 2
At1g08990 ----------
At4g33330 ----------MTIMTMI 7
At1g77130 ----------MIPS 4
At3g18660 ----------MANSPAAPAPTT 12
At1g56600 ----------
At1g09350 ----------
At2g47180 ----------
At1g60450 ----------
At1g60470 ----------
At4g26250 ----------
At5g23790 ----------


At3g61130 LDWRERLAMQSVRSLFSKEILDVIATSTADLGPLSLDSFKKNNLSASWRGTGVDPSFRHS 120
At2g46480 ----------MTDACCLKGNEDKMVPRFGHGTWIGKAFN-- 29
At4g38270 NIRLGRVKMGDLSPSWVLENPAYQVSRKTKGSQLVIPRDSFQNDTGMEDNASHSTTNQTD 116
At5g47780 LNLLPRE----------SPAVLRGGLVGAVYSDKNSRRLDQLSARVLSATD- 95
At2g30575 R----------LTAIEQDE-EGLKEPKRILQ-DRDFNSVVLSNSSD 91
At1g06780 R----------LSAIEHEDGEGLKGPRLILFKDGEFNSSAES---- 88
At2g38650 SPGFVTVQPASSFESFTRINATKHTQRDVSERVDEVLQKINPVLPKKSDINVGSRDVNAT 108
At3g25140 ----------MANHHRLLRGGGSPAII- 17
At3g02350 ----------MAVAFRGGRGGVGSGQS- 17
At2g20810 ----------MRRRGGDSFRR- 11
At1g18580 ----------MRRWPVDHRR- 10
At5g54690
At3g01040
At5g15470 ----------MRSGRRPQGRRIAIRNETETELRSRIGEKSGSEKRNL 37
At3g58790
At1g19300
At3g50760
At1g13250
At3g06260
At1g02720
At4g02130
At3g62660
At1g24170
At1g70090
At3g28340
At1g54940 TKTHNSRGKIFMIYLILVSLSLLGLILP----------FKP---- 33
At1g08990
At4g33330 MKMAPSKSALIRFNLVLLGFSFLLYTAI----------FFHPSSS 42
At1g77130 SSPMESRHRLSFSNEKTSRRRFQRIEKG----------VKFNT 37
At3g18660 TTGGDSRRRLSASIEAICKRRFRRNSKGGGRSDMVKPFNIINFSTQDKNSSCCCFTKFQI 72
At1g56600
At1g09350
At2g47180
At1g60450
At1g60470
At4g26250
At5g23790


At3g61130 ENPATPDVKSNNLNEKRDSISKDSIHQKVETPTKIHRRQLREKRREMRANELVQHNDDTI 180
At2g46480 ----------DTPEMLHERSLRQEKRLERANELMN--DDSL 58
At4g38270 ES----------ENQFPNVDFASPAKLKRQILRQERRGQRTLELIRQEKETD 158
At5g47780 ----------DDTHSHTDISIKQVTHDAASDSHINRENMHV 126
At2g30575 KS----------NDTVQSNEGDQKNFLSEVDKGNNHKPKEE-QAVSQKTTVSSNAE 136
At1g06780 ----------DGGNTYKNREEQVIVSQKMTVSSDEK 114
At2g38650 SG----------TDSKKRGLPVSPTVVANPSPANKTKSEASYTG 142
At3g25140 ----------GGR---ITLTAFASTIALFLFTLSFFFASDS 45
At3g02350 ----------TGLRSFFSYRIFISALFSFLFLATFSVVLNS 48
At2g20810 ----------AGRRKISNVVWWVLS--GIALLLFFLILSKA 40
At1g18580 ----------RGRRRLSSWIWFLLGSFSVAGLVLFIVQHYH 41
At5g54690 ----------MQLHISPS-LRHVTVVT-------GKG 19
At3g01040 ----------
At5g15470 IG----------SDVGMQLHISPS-MRSITISS-------SNE 62
At3g58790 ----------MKFYISATGIKKVTISNPGVGIGKGSG 27
At1g19300 ----------
At3g50760 ----------
At1g13250 ----------
At3g06260 ----------
At1g02720 ----------
At4g02130 ----------
At3g62660 ----------
At1g24170 ----------
At1g70090 ----------
At3g28340 ----------
At1g54940 ----------LFRITSPSSTLRIDLPSPQVNKNPKWLRLIRNYL 67
At1g08990 ----------
At4g33330 VYFSSGASFVGCSFRDCTPKVVRGVKMQELVEENEINKKDLLTASNQTKLEAPSFMEEIL 102
At1g77130 LKLVLICIMLGALFTIYRFRYPPLQIPEIPTSFGLTTDPRYVATAEINWNHMSNLVEKHV 97
At3g18660 VKLLLFILLSATLFTIIYS--PEAYHHSLSHSSSSRRQDPRYFSDLDINWDDVTKTLENIE 130
At1g56600 ----------
At1g09350 ----------
At2g47180 ----------
At1g60450 ----------
At1g60470 ----------
At4g26250 ----------
At5g23790 ----------

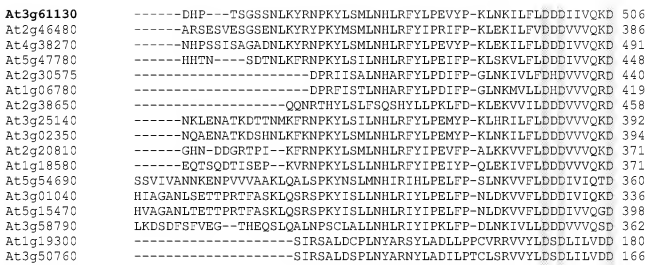
```
At3g61130       LKLENAAIERSKSVDSAVLGKYSIWRRENENDNS-DSNIRLMRDQVIMARVYSGIAK--- 236
At2g46480       QKLETAAMARSRSVDSAPLGNYTIWKNEYRRGKSFEDMLRLMQDQIIMARVYSGLAK--- 115
At4g38270       EQMQEAAIQKSMSFENSVIGKYSIWRRDYESPNA-DAILKLMRDQIIMAKAYANIAK--- 214
At5g47780       QLTQQTSEKVDEQPEPNAFG----AKKDTGNVLMPDAQVRHLKDQLIRAKVYLSLPS--- 179
At2g30575       VKISARDIQLNHKTEFRPPSSKSEKNTRVQLERATDERVKEIRDKIIQAKAYLNLAL--- 193
At1g06780       GQILPTVNQLANKTDFKPPLSKGEKNTRVQPDRATDVKTKEIRDKIIQAKAYLNFAP--- 171
At2g38650       VQRKIVSGDETWRTCEVKYGSYCLWREENKEPMKDAKVKQMKDQLFVARAYYPSIAK--- 199
At3g25140       NDSPDLLLPGVEYSNGVGSRRSMLDIKSDPLKPRLIQIRKQADDHRSLALAYASYARKLK 105
At3g02350       SRHQPHQDHTLPSMGNAYMQRTFLALQSDPLKTRLDLIHKQAIDHLTLVNAYAAYARKLK 108
At2g20810       GHIEPRPSIPKRRYRNDKFVEG-MNMTEEMLSP--TSVARQVNDQIALAKAFVVIAK--- 94
At1g18580       HQQDPSQLLLERDTRTEMVSPPHLNFTEEVTSA--SSFSRQLAEQMTLAKAYVFIAK--- 96
At5g54690       LREFIKVKVGSRRFSYQMVFYSLLFFTFLLRFVFVLSTVDTIDGDPSPCSSLACLGKRLK 79
At3g01040       ----MKIKVAARHISYRTLFHTILILAFLLPFVFILTAVVTLEG-VNKCSSFDCFGRRLG 55
At5g15470       FIDLMKIKVAARHISYRTLFHTILILAFLLPFVFILTAVVTLEG-VNKCS---SIGRRIG 118
At3g58790       GCAAAAAALAARRFSSRTLLLLLLLLAIVLPFIFVRFAFLVLESASVCDSPLDCMGLRLF 87
At1g19300       ------------------------------------------------------------
At3g50760       ------------------------------------------------------------
At1g13250       ------------------------------------------------------------
At3g06260       ------------------------------------------------------------
At1g02720       ------------------------------------------------------------
At4g02130       ------------------------------------------------------------
At3g62660       ------------------------------------------------------------
At1g24170       ------------------------------------------------------------
At1g70090       ------------------------------------------------------------
At3g28340       ------------------------------------------------------------
At1g54940       PE--KRIQVGFLNIDEKERESYEARGPL-----VLKNIHVPLDHIPKNVTWKSLYPEWIN 120
At1g08990       ------------------------------------------------------------
At4g33330       TRGLGKTKIGMVNMEECDLTNWKRYG---------ETVHIHFERVSKLFKWQDLFPEWID 153
At1g77130       FGRSEYQGIGLINLNDNEIDRFKEVTKS-----DCDHVALHLDYAAKNITWESLYPEWID 152
At3g18660       EGRT----IGVLNFDSNEIQRWREVSKSKDNGDEEKVVVLNLDYADKNVTWDALYPEWID 186
At1g56600       ------------------------------------------------------------
At1g09350       ------------------------------------------------------------
At2g47180       ------------------------------------------------------------
At1g60450       ------------------------------------------------------------
At1g60470       ------------------------------------------------------------
At4g26250       ------------------------------------------------------------
At5g23790       ------------------------------------------------------------
```


At3g61130 --LKNKNDLLQELQARLKDSQ----------RVLGEATSDADLPRSAHEKLRAMGQVLA 283
At2g46480 --FTNNLALHQEIETQLMK------------LAWEEESTDIDQEQRVLDSIRDMGQILA 160
At4g38270 --SKNVTNLYVFLMQQCGENK----------RVIGKATSDADLPSSALDQAKAMGHALS 261
At5g47780 --AKANAHFVRELRLRIKEVQ----------RALADASKDSDLPKTAIEKLKAMEQTLA 226
At2g30575 --PGNNSQIVKELRVRTKELER----------ATGDTTKDKYLPKSSPNRLKAMEVALY 240
At1g06780 --PGSNSQVVKELRGRLKELER----------SVGDATKDKDLSKGALRRVKPMENVLY 218
At2g38650 --MPSQSKLTRDMKQNIQEFER----------ILSESSQDADLPPQVDKKLQKMEAVIA 246
At3g25140 LENSKLVRIFADLSRNYTDLINKPTYRALYDSDGASIEESVLRQFEKEVKERIKMTRQVI 165
At3g02350 LDASKQLKLFEDLAINFSDLQSKPGLKSAVSDNGNALEEDSFRQLEKEVKDKVKTARMMI 168
At2g20810 --ESKNLQFAWDLSAQIRNSQ------LLLSSAATRRSPLTVLESESTIRD---MAVLLY 143
At1g18580 --EHNNLHLAWELSSKIRSCQ------LLLSKAAMRGQPISFDEAKPIITG---LSALIY 145
At5g54690 PKLLGRRVDSGNVPEAMY----------QVLEQPLSEQELKGRSDIPQTLQDFMS 124
At3g01040 PRLLGRIDDSEQRLVRDF----------YKILNEVSTQEIPDGLKLPESFSQLVS 100
At5g15470 PRLLGRVDDSE-RLARDF----------YKILNEVSTQEIPDGLKLPNSFSQLVS 162
At3g58790 R----GGDTSLKIGEELT----------RALVEETTDHQDVNGRGTKGSLESFDD 128
At1g19300 ----------MSQHLLLLILLSL 13
At3g50760 ----------MHSKFILYLSI 11
At1g13250 ----------MSS--------LRLRLCL 10
At3g06260 ----------MASRSLSYTQLLGLLSFI 18
At1g02720 ----------MHWITRFSAFFSAALAM 17
At4g02130 ----------MLWITRFAGLFSAAMAV 17
At3g62660 ----------MRFSGLFSAALVI 13
At1g24170 ----------MSSRFSLTVVC 11
At1g70090 ----------MRLRFPMKSAVLAFAIF 17
At3g28340 ----------MMSGSRLASRLIIIFSI 17
At1g54940 ----EEASTCPEIPLPQPEGSDANVDVIVARVPCDGWSANKGLRDVFRLQVNLAAANLAV 176
At1g08990 ----------MEGSEADVDVVVVKVPCDGFSEKRGLRDVFRLQVNLAAANLVV 43
At4g33330 EEEETEVPTCPEIPMPDFESLEK-LDLVVVKLPCN-YPEEGWRREVLRLQVNLVAANLAA 211
At1g77130 EVEEFEVPTCPSLPLIQ-IPGKPRIDLVIAKLPCD--KSGKWSRDVARLHLQLAAARVAA 209
At3g18660 EEQETEVPVCPNIPNIK-VPTR-RLDLIVVKLPCR--KEGNWSRDVGRLHLQLAAATVAA 242
At1g56600 ----------
At1g09350 ----------
At2g47180 ----------
At1g60450 ----------
At1g60470 ----------
At4g26250 ----------
At5g23790 ----------

```
At3g61130      KAKMQLYDCKLVTGKLRAMLQTADEQVRSLKKQSTFLAQLAAKTIPNPIHCLSMRLTIDY 343
At2g46480      RAHEQLYECKLVTNKLRAMLQTVEDELENEQTYITFLTQLASKALPDAIHCLTMRLNLEY 220
At4g38270      LAKDELYDCHELAKKFRAILQSTERKVDGLKKKGTFLIQLAAKTFPKPLHCLSLQLAADY 321
At5g47780      KGKQIQDDCSTVVKKLRAMLHSADEQLRVHKKQTMFLTQLTAKTIPKGLHCLPLRLTTDY 286
At2g30575      KVSRAFHNCPAIATKLQAMTYKTEEQARAQKKQAAYLMQLAARTTPKGLHCLSMRLTTEY 300
At1g06780      KASRVFNNCPAIATKLRAMNYNTEEQVQAQKNQAAYLMQLAARTTPKGLHCLSMRLTSEY 278
At2g38650      KAKSFPVDCNNVDKKLRQILDLTEDEASFHMKQSVFLYQLAVQTMPKSLHCLSMRLTVEH 306
At3g25140      AEAKESFDNQLKIQKLKDTIFAVNEQLTNAKKQGAFSSLIAAKSIPKGLHCLAMRLMEER 225
At3g02350      VESKESYDTQLKIQKLKDTIFAVQEQLTKAKKNGAVASLISAKSVPKSLHCLAMRLVGER 228
At2g20810      QAQQLHYDSATMIMRLKASIQALEEQMSSVSEKSSKYGQIAAEEVPKSLYCLGVRLTTEW 203
At1g18580      KAQDAHYDIATTMMTMKSHIQALEERANAATVQTTIFGQLVAEALPKSLHCLTIKLTSDW 205
At5g54690      EVKRSKSDAREFAQKLKEMVTLMEQRTRTAKIQEYLYRHVASSSIPKQLHCLALKLANEH 184
At3g01040      DMKNNHYDAKTFALVFRAMVEKFERDLRESKFAELMNKHFAASSIPKGIHCLSLRLTDEY 160
At5g15470      DMKNNHYDAKTFALVLRAMMEKFERDMRESKFAELMNKHFAASSIPKGIHCLSLRLTDEY 222
At3g58790      LVKEMTLKRRDIRAFASVTKKMMERKVQSAKHHELVYWHLASHGIPKSLHCLSLRLTEEY 188
At1g19300      LLLHKPISATT----------------------IIQKFKEAPQFYNSADCPLIDDSESD 50
At3g50760      LAVFTVSFAG-----------------------GERFKEAPKFFNSPECLTIENDED- 45
At1g13250      LLLLPITISCVTV-------TLTDLP-----------AFREAPAFRNGRECSKTTWIPSD 52
At3g06260      LLLVTTTTMAVRVGVILHKPSAPTLP-----------VFREAPAFRNGDQCG-------- 59
At1g02720      ILLSPSLQSFSPAAAIRSSHPYADEFKPQQNS--DYSSFRESPMFRNAEQCRSSGEDSGV 75
At4g02130      IVLSPSLQSFPPAAAIRSSP--SPIFR-----------KAPAVFNNGDECLSSG---GV 60
At3g62660      IVLSPSLQSFPPAEAIRSSHLDAYLRFPSSDPPPHRFSFRKAPVFRNAADCAAADIDSGV 73
At1g24170      LIALLPFVVGIRLIPARITSVGDGGGGGGNNGFSKLGPFMEAPEYRNGKECVSSSVNREN 71
At1g70090      LVFIPLFSVGIRMIPGRLTAVSATVGNG-----FDLGSFVEAPEYRNGKECVSQSLNREN 72
At3g28340      ISTSFFTVESIRLFPDSFDDASS--------------DLMEAPAYQNGLDCSVLAKNRLL 63
At1g54940      QSGLRTVNQAVYVVFIGSCGPMHEIFPCDERVMRVEDYWVYKPYLPRLKQKLLMPVGSCQ 236
At1g08990      ESGRRNVDRTVYVVFIGSCGPMHEIFRCDERVKRVGDYWVYRPDLTRLKQKLLMPPGSCQ 103
At4g33330      KKGKTDWRWKSKVLFWSKCQPMIEIFRCDDLEKREADWWLYRPEVVRLQQRLSLPVGSCN 271
At1g77130      SS---KGLHNVHVILVSDCFPIPNLFTGQELVARQGNIWLYKPNLHQLRQKLQLPVGSCE 266
At3g18660      SA---KGFFRGHVFFVSRCFPIPNLFRCKDLVSRRGDVWLYKPNLDTLRDKLQLPVGSCE 299
At1g56600      ------------------------------------------------MAPEINT-- 7
At1g09350      ------------------------------------------------MAPEMNN-- 7
At2g47180      ------------------------------------------------MAPGLTQTA 9
At1g60450      ----------------------------------------------MTPETHVD- 8
At1g60470      ----------------------------------------------MAPEISVN- 8
At4g26250      --------------------------------------------MAQMSMTVEKS- 11
At5g23790      --------------------------------------------MTMTVEKR- 8
```

```
At3g61130      YLLSP----EKRKFPRSENLENPNLYHYALFSDNVLAASVVVNSTIMNAKDPSKHVFHLV 399
At2g46480      HLLPL----PMRNFPRRENLENPKLYHYALFSDNVLAASVVVNSTVMNAQDPSRHVFHLV 276
At4g38270      FILGFNEEDAVKEDVSQKKLEDPSLYHYAIFSDNVLATSVVVNSTVLNAKEPQRHVFHIV 381
At5g47780      YALNS----SEQQFPNQEKLEDTQLYHYALFSDNVLATSVVVNSTITNAKHPLKHVFHIV 342
At2g30575      FTLDH----EKRQLL-QQSYNDPDLYHYVVFSDNVLASSVVVNSTISSSKEPDKIVFHVV 355
At1g06780      FSLDP----EKRQMPNQQNYFDANFNHYVVFSDNVLASSVVVNSTISSSKEPERIVFHVV 334
At2g38650      FKSDS------LEDPISEKFSDPSLLHFVIISDNILASSVVVINSTVVHARDSKNFVFHVL 360
At3g25140      IAHP--EKYTDEGKDRPRELEDPNLYHYAIFSDNVIAASVVVNSAVKNAKEPWKHVFHVV 283
At3g02350      ISNP--EKYKDAPPDP--AAEDPTLYHYAIFSDNVIAVSVVVRSVVMNAEEPWKHVFHVV 284
At2g20810      FQNLDLQRTLKERSRVDSKLTDNSLYHFCVFSDNIIATSVVVNSTALNSKAPEKVVFHLV 263
At1g18580      VTEP-SRHELADENRNSPRLVDNNLYHFCIFSDNVIATSVVVNSTVSNADHPKQLVFHHV 264
At5g54690      SINAAAR-LQLPEAELVPMLVDNNYFHFVLASDNILAASVVAKSLVQNALRPHKIVLHII 243
At3g01040      SSNAHAR-RQLPSPELLPVLSDNAYHHFVLATDNILAASVVVSSAVQSSSKPEKIVFHVI 219
At5g15470      SSNAHAR-RQLPSPEFLPVLSDNAYHHFILSTDNILAASVVVSSAVQSSSKPEKIVFHII 281
At3g58790      SVNAMAR-MRLPPPESVSRLTDPSFHHIVLLTDNVLAASVVISSTVQNAVNPEKFVFHVV 247
At1g19300      DDVVAKP---------IFCSRRAVHVAMTLDAAYIRGSVAAVLSVLQHSSCPENIVFHFV 101
At3g50760      ----------------FVCSDKAIHVAMTLDTAYLRGSMAVILSVLQHSSCPQNIVFHFV 89
At1g13250      ----------------HEHNPSIIHIAMTLDAIYLRGSVAGVFSVLQHASCPENIVFHFI 96
At3g06260      ----------------TREADQIHIAMTLDTNYLRGTMAAVLSLLQHSTCPENLSFHFL 102
At1g02720      CN----------------PNLVHVAITLDIDYLRGSIAAVNSILQHSMCPQSVFFHFL 117
At4g02130      CN----------------PSLVHVAITLDVEYLRGSIAAVNSILQHSVCPESVFFHFI 102
At3g62660      CN----------------PSLVHVAITLDFEYLRGSIAAVHSILKHSSCPESVFFHFL 115
At1g24170      -FVSSSS---------SSNDPSLVHIAMTLDSEYLRGSIAAVHSVLRHASCPENVFFHFI 121
At1g70090      -FVSS-------------CDASLVHVAMTLDSEYLRGSIAAVHSMLRHASCPENVFFHLI 118
At3g28340      -LACD---------------PSAVHIAMTLDPAYLRGTVSAVHSILKHTSCPENIFFHFI 107
At1g54940      IAPSFAQ-FGQEAWRPKHEDNLASKAVTALPRRLRVAYVTVLHSSEAYVCGAIALAQSIR 295
At1g08990      IAP--LG-QGEAWIQDKNRNLTSEKTTLSSFTAQRVAYVTLLHSSEVYVCGAIALAQSIR 160
At4g33330      LALP------LWAPQGVDKVYDLTKIEAETKRPKREAYVTVLHSSESYVCGAITLAQSLL 325
At1g77130      LSVP----------------LQAKDNFYSAGAKKEAYATILHSAQFYVCGAIAAAQSIR 309
At3g18660      LSLP----------------LGIQDRPSLGNPKREAYATILHSAHVYVCGAIAAAQSIR 342
At1g56600      --------------------KLTVPVHSATGGEKRAYVTFLAGTGDYVKGVVGLAKGLR 46
At1g09350      --------------------KLSY------GEKKRAYVTFLAGTGDYVKGVVGLAKGLR 40
At2g47180      DAMS---------------TVTITKPSLPSVQDSDRAYVTFLAGNGDYVKGVVGLAKGLR 54
At1g60450      ---------------------MINASEKAPKE---RAYVTFLAGNGDYVKGVVGLAKGLR 44
At1g60470      ---------------------PMYLSEKAHQAPPRRAYVTFLAGNGDYVKGVVGLAKGLR 47
At4g26250      ----------------------IKADVTVSHDRVKRAYVTFLAGNKDYWMGVVGLAKGLR 49
At5g23790      ----------------------IEADVTVSHEGVERAYVTFLAGNKDYWMLVVGLAKGLR 46
                                                    .         .          .
```

```
At3g61130       TDKLNFGAMNMWFLLN--PPGKATIHVENVDEFKWLNSSYCPVLRQLESAAMREYYFKA- 456
At2g46480       TDKLNFGAMSMWFLLN--PPGEATIHVQRFEDFTWLNSSYSPVLSQLESAAMKKFYFKT- 333
At4g38270       TDKLNFGAMKMWFRIN--APADATIQVENINDFKWLNSSYCSVLRQLESARLKEYYFKA- 438
At5g47780       TDRLNYAAMRMWFLDN--PPGKATIQVQNVEEFTWLNSSYSPVLKQLSSRSMIDYYFRA- 399
At2g30575       TDSLNYPAISMWFLLN--PSGRASIQILNIDEMNVLPLYHAELLMKQNSS---------- 403
At1g06780       TDSLNYPAISMWFLLN--IQSKATIQILNIDDMDVLPRDYDQLLMKQNSN---------- 382
At2g38650       TDEQNYFAMKQWFIRN--PCKQSTVQVLNIEKLELDDSDMKLSLSAEFRVSFPSGDLLAS 418
At3g25140       TDKMNLGAMQVMFKLK-EYK-GAHVEVKAVEDYTFLNSSYVPVLKQLESANLQKFYFE-- 339
At3g02350       TDRMNLAAMKVWFKMR-PLDRGAHVEIKSVEDFKFLNSSYAPVLRQLESAKLQKFYFE-- 341
At2g20810       TNEINYAAMKAWFAINMDNLRGVTVEVQKFEDFSWLNASYVPVLKQLQDSDTQSYYFS-- 321
At1g18580       TNRVSYKAMQAWFLS--NDFKGSAIEIRSVEEFSWLNASYSPVVKQLLDTDARAYYFG-- 320
At5g54690       TDRKTYFPMQAWFSLH--PLSPAIIEVKALHHFDWLSKGKVPVLEAMEKDQRVRSQFRGG 301
At3g01040       TDKKTYAGMHSWFALN--SVAPAIVEVKSVHQFDWLTRENVPVLEAVESHNSIRNYYHGN 277
At5g15470       TDKKTYAGMHSWFALN--SVAPAIVEVKGVHQFDWLTRENVPVLEAVESHNGVRDYYHGN 339
At3g58790       TDKKTYTPMHAWFAIN--SASSPVVEVKGLHQYDWPQEVNFKVREMLDIHRLIWRRHYQN 305
At1g19300       ASASADASS--LRATIS----SSFPYLD-FTVYVFNVSSVSRLISS-------------- 140
At3g50760       TSK--QSHR--LQNYVV----ASFPYLK-FRIYPYDVAAISGLIST-------------- 126
At1g13250       ATHRRSAD---LRRIIS----STFPYLT-YHIYHFDPNLVRSKISS-------------- 134
At3g06260       SLPHTEND---LFTSIK----STFPYLN-FKIYQFDPNLVRSKISK-------------- 140
At1g02720       VSSESQN----LESLIR----STFPKLTNLKIYYFAPETVQSLISS-------------- 155
At4g02130       AVSEETNL---LESLVR----SVFPRLK-FNIYDFAPETVRGLISS-------------- 140
At3g62660       VS---ETD---LESLIR----STFPELK-LKVYYFDPEIVRTLIST-------------- 150
At1g24170       AAEFDSASPRVLSQLVR----STFPSLN-FKVYIFREDTVINLISS-------------- 162
At1g70090       AAEFDPASPRVLSQLVR----STFPSLN-FKVYIFREDTVINLISS-------------- 159
At3g28340       ASGTSQGS---LAKTLS----SVFPSLS-FKVYTFEETTVKNLISS-------------- 145
At1g54940       QSGSHKDMILLHDHTIT------NKSLIGLSAAGWNLRLIDRIRSP-------------- 335
At1g08990       QSGSTKDMILLHDDSIT------NISLIGLSLAGWKLRRVERIRSP-------------- 200
At4g33330       QTNTKRDLILLHDDSIS------ITKLRALAAAGWKLRRIIRINP--------------- 365
At1g77130       MSGSTRDLVILVDETIS------EYHKSGLVAAGWKIQMFQRIRNP-------------- 349
At3g18660       QSGSTRDLVILVDDNIS------GYHRSGLEAAGWQIRTIQRIRNP-------------- 382
At1g56600       KAKSKYPLVVAVLPDVP------EDHRKQLVDQGCVVKEIEPVYPP-------------- 86
At1g09350       KTKSKYPLVVAVLPDVP------ADHRRQLLDQGCVIKEIQPVYPP-------------- 80
At2g47180       KVKSAYPLVVAMLPDVP------EEHRRILVDQGCIVREIEPVYPP-------------- 94
At1g60450       KVKSAYPLVVAMLPDVP------EEHREILRSQGCIVREIEPVHPP-------------- 84
At1g60470       KVKSAYPLVVAMLPDVP------EEHREILRSQGCVVREIEPVYPP-------------- 87
At4g26250       KVKSAYPLVVAILPDVP------EEHRQILLAQGCIIREIEPVYPP-------------- 89
At5g23790       KVKSAYPLVVATLPDVP------EEHRQILVDQGCIIRDIEPVYPP-------------- 86
```


At3g61130 ------DHP---TSGSSNLKYRNPKYLSMLNHLRFYLPEVYP-KLNKILFLDDDIIVQKD 506
At2g46480 ------ARSESVESGSENLKYRYPKYMSMLNHLRFYIPRIFP-KLEKILFVDDDVVVQKD 386
At4g38270 ------NHPSSISAGADNLKYRNPKYLSMLNHLRFYLPEVYP-KLEKILFLDDDIVVQKD 491
At5g47780 ------HHTN----SDTNLKFRNPKYLSILNHLRFYLPEIFP-KLSKVLFLDDDIVVQKD 448
At2g30575 ----------------------DPRIISALNHARFYLPDIFP-GLNKIVLFDHDVVVQRD 440
At1g06780 ----------------------DPRFISTLNHARFYLPDIFP-GLNKMVLLDHDVVVQRD 419
At2g38650 ----------------------QQNRTHYLSLFSQSHYLLPKLFD-KLEKVVILDDDVVVQRD 458
At3g25140 ------NKLENATKDTTNMKFRNPKYLSILNHLRFYLPEMYP-KLHRILFLDDDVVVQKD 392
At3g02350 ------NQAENATKDSHNNLKFKNPKYLSMLNHLRFYLPEMYP-KLNKILFLDDDVVVQKD 394
At2g20810 ------GHN-DDGRTPI--KFRNPKYLSMLNHLRFYIPEVFP-ALKKVVFLDDDVVVQKD 371
At1g18580 ------EQTSQDTISEP--KVRNPKYLSLLNHLRFYIPEIYP-QLEKIVFLDDDVVVQKD 371
At5g54690 SSVIVANNKENPVVVAAKLQALSPKYNSLMNHIRIHLPELFP-SLNKVVFLDDDIVIQTD 360
At3g01040 HIAGANLSETTPRTFASKLQSRSPKYISLLNHLRIYLPELFP-NLDKVVFLDDDIVIQKD 336
At5g15470 HVAGANLTETTPRTFASKLQSRSPKYISLLNHLRIYIPELFP-NLDKVVFLDDDIVVQGD 398
At3g58790 LKDSDFSFVEG--THEQSLQALNPSCLALLNHLRIYIPKLFP-DLNKIVLLDDDVVVQSD 362
At1g19300 ----------------------SIRSALDCPLNYARSYLADLLPPCVRRVVYLDSDLILVDD 180
At3g50760 ----------------------SIRSALDSPLNYARNYLADILPTCLSRVVYLDSDLILVDD 166
At1g13250 ----------------------SIRRALDQPLNYARIYLADLLPIAVRRVIYFDSDLVVVDD 174
At3g06260 ----------------------SIRQALDQPLNYARIYLADIIPSSVDRIIYLDSDLVVVDD 180
At1g02720 ----------------------SVRQALEQPLNYARNYLADLLEPCVKRVIYLDSDLVVVDD 195
At4g02130 ----------------------SVRQALEQPLNYARSYLADLLEPCVNRVIYLDSDLVVVDD 180
At3g62660 ----------------------SVRQALEQPLNYARNYLADLLEPCVRRVIYLDSDLIVVDD 190
At1g24170 ----------------------SIRLALENPLNYARNYLGDILDRSVERVIYLDSDVITVDD 202
At1g70090 ----------------------SIRQALENPLNYARNYLGDILDPCVDRVIYLDSDIIVVDD 199
At3g28340 ----------------------SIRQALDSPLNYARSYLSEILSSCVSRVIYLDSDVIVVDD 185
At1g54940 ----------------------FSQ---KDSYNEWNYSKLRVWQVTDYDKLVFIDADFIILKK 373
At1g08990 ----------------------FSK---KRSYNEWNYSKLRVWQVTDYDKLVFIDADFIIVKN 238
At4g33330 ----------------------LAE---KDSYNEYNYSKFRLWQLTDYDKVIFIDADIIVLRN 403
At1g77130 ----------------------NAV---PNAYNEWNYSKFRLWQLTEYSKIIFIDADMLILRN 387
At3g18660 ----------------------KAE---KDAYNEWNYSKFRLWQLTDYDKIIFIDADLLILRN 420
At1g56600 ----------------------ENQTEFAMAYYVINYSKLRIWEFVEYNKMIYLDGDIQVFDN 127
At1g09350 ----------------------DNQTQFAMAYYVLNYSKLRIWKFVEYSKLIYLDGDIQVFEN 121
At2g47180 ----------------------ENQTQFAMAYYVINYSKLRIWKFVEYSKMIYLDGDIQVYEN 135
At1g60450 ----------------------DSQDAYARAYYIINYSKLRIWNFEEYNKMIYLDADIQVFGN 125
At1g60470 ----------------------DNQVEFAMAYYVLNYSKLRIWNFEEYSKMIYLDADIQVFDN 128
At4g26250 ----------------------ENKTGYSMAYYVINYSKLRIWEFVEYEKMIYLDGDIQVFSN 130
At5g23790 ----------------------ENTTGYSMAYYVINYSKLRIWEFVEYEKMIYLDGDIQVFKN 127
: ::: .* *. .

```
At3g61130      LTPLWEVNLNGKVNGAVETCGE-----SFHRFDKYLNFSNPHIARNFNPN----ACGWAY 557
At2g46480      LTPLWSIDLKGKVN-------------------------------ENFDPK----FCGWAY 412
At4g38270      LAPLWEIDMQGKVNGAVETCKE-----SFHRFDKYLNFSNPKISENFDAG----ACGWAF 542
At5g47780      LSGLWSVDLKGNVNGAVETCGE-----SFHRFDRYLNFSNPLISKNFDPR----ACGWAY 499
At2g30575      LTRLWSLDMTGKVVGAVETCLEGD--PSYRSMDSFINFSDAWVSQKFDPK----ACTWAF 494
At1g06780      LSRLWSIDMKGKVVGAVETCLEGE--SSFRSMSTFINFSDTWVAGKFSPR----ACTWAF 473
At2g38650      LSPLWDLDMEGKVNGAVKSCTVR------------LGQLRSLKRGNFDTN----ACLWMS 502
At3g25140      LTGLWEIDMDGKVNGAVETCFG-----SFHRYAQYMNFSHPLIKEKFNPK----ACAWAY 443
At3g02350      VTGLWKINLDGKVNGAVETCFG-----SFHRYGQYLNFSHPLIKENFNPS----ACAWAF 445
At2g20810      LSSLFSIDLNKNVNGAVETCME-----TFHRYHKYLNYSHPLIRSHFDPD----ACGWAF 422
At1g18580      LTPLFSLDLHGNVNGAVETCLE-----AFHRYYKYLNFSNPLISSKFDPQ----ACGWAF 422
At5g54690      LSPLWDIDMNGKVNGAVETCRGEDKFVMSKKFKSYLNFSNPTIAKNFNPE----ECAWAY 416
At3g01040      LSPLWDIDLNGKVNGAVETCRGEDVWVMSKRLRNYFNFSHPLIAKHLDPE----ECAWAY 392
At5g15470      LTPLWDVDLGGKVNGAVETCRGEDEWVMSKRLRNYFNFSHPLIAKHLDPE----ECAWAY 454
At3g58790      LSSLWETDLNGKVVGAVVDSWCGDNCCPGRKYKDYFNFSHPLISSNLVQE----DCAWLS 418
At1g19300      IAKLAATDLGRDSVLAAPEYCNAN----FTSYFTSTFWSNPTLSLTFADR---KACYFNT 233
At3g50760      ISKLFSTHIPTDVVLAAPEYCNAN----FTTYFTPTFWSNPSLSITLSLNRRATPCYFNT 222
At1g13250      VAKLWRIDLR-RHVVGAPEYCHAN----FTNYFTSRFWSSQGYKSALKDR---KPCYFNT 226
At3g06260      IEKLWHVEME-GKVVAAPEYCHAN----FTHYFTRTFWSDPVLVKVLEGK---RPCYFNT 232
At1g02720      IVKLWKTGLG-QRTIGAPEYCHAN----FTKYFTGGFWSDKRFNGTFKGR---NPCYFNT 247
At4g02130      IAKLWKTSLG-SRIIGAPEYCHAN----FTKYFTGGFWSEERFSGTFRGR---KPCYFNT 232
At3g62660      IAKLWMTKLG-SKTIGAPEYCHAN----FTKYFTPAFWSDERFSGAFSGR---KPCYFNT 242
At1g24170      ITKLWNTVLTGSRVIGAPEYCHAN----FTQYFTSGFWSDPALPGLISGQ---KPCYFNT 255
At1g70090      ITKLWNTSLTGSRIIGAPEYCHAN----FTKYFTSGFWSDPALPGFFSGR---KPCYFNT 252
At3g28340      IQKLWKISLSGSRTIGAPEYCHAN----FTKYFTDSFWSDQKLSSVFDSK---TPCYFNT 238
At1g54940      LDHLFYYPQLSASGNDKVLFNSGIMVLEPSACMFKDLMEKSFKIESYNGG---DQG--FL 428
At1g08990      IDYLFSYPQLSAAGNNKVLFNSGVMVLEPSACLFEDLMLKSFKIGSYNGG---DQG--FL 293
At4g33330      LDLLFHFPQMSATGNDVWIYNSGIMVIEPSNCTFTTIMSQRSEIVSYNGG---DQG--YL 458
At1g77130      IDFLFEFPEISATGNNATLFNSGLMVVEPSNSTFQLLMDNINEVVSYNGG---DQG--YL 442
At3g18660      IDFLFSMPEISATGNNGTLFNSGVMVIEPCNCTFQLLMEHINEIESYNGG---DQG--YL 475
At1g56600      IDHLFDLPNGQFYAVMDCFCEKTWSHSPQYKIGYCQQCPDKVTWPEAKLG---PKPPLYF 184
At1g09350      IDHLFDLPDGNFYAVKDCFCEKTWSHTPQYKIGYCQQCPDKVTWPESELG---PKPPLYF 178
At2g47180      IDHLFDLPDGYLYAVMDCFCEKTWSHTPQYKIRYCQQCPDKVQWPKAELG---EPPALYF 192
At1g60450      IDDLFDMQDGYLHGVLSCFCEKIWSYTPLYSIGYCQYCPEKVVWPAEMES---APPSPYF 182
At1g60470      IDHLFDLSDAYFYAVMDCFCEKTWSHSLQYSIGYCQQCPEKVTWPEDMES---PPPLYF 185
At4g26250      IDHLFDTPRGYLYAVKDCFCEISWSKTPQFKIGYCQQCPEKVTWPVESLG---SPPPVYF 187
At5g23790      IDHLFDTPRGYLYAVKDCFCEVSWSKTPQYKIGYCQQSPEKVTWPVESLG---APPPVYF 184
               :  *
```

```
At3g61130     GMNMFDLKEWKKRDITGIYHKWQNMNENR--TLWKLGTLPPGLITFYGLTHPLNKAWHVL 615
At2g46480     GMNIFDLKEWKKNNITETYHFWQNLNENR--TLWKLGTLPPGLITFYNLTQPLQRKWHLL 470
At4g38270     GMNMFDLKEWRKRNITGIYHYWQDLNEDR--TLWKLGSLPPGLITFYNLTYAMDRSWHVL 600
At5g47780     GMNVFDLDEWKRQNITEVYHRWQDLNQDR--ELWKLGTLPPGLITFWRRTYPLDRKWHIL 557
At2g30575     GMNLFDLEEEWRRQELTSVYLKYFDLGVKG--HLWKAGGLPVGWLTFFGQTFPLEKRWNVG 552
At1g06780     GMNLIDLEEWRIRKLTSTYIKYFNLGTKR--PLWKAGSLPIGWLTFYRQTLALDKRWHVM 531
At2g38650     GLNVVDLARWRALGVSETYQKYYKEMSSGD-ESSEAIALQASLLTFQDQVYALDDKWALS 561
At3g25140     GMNFFDLDAWRREKCTEEYHYWQNLNENR--ALWKLGTLPPGLITFYSTTKPLDKSWHVL 501
At3g02350     GMNIFDLNAWRREKCTDQYHYWQNLNEDR--TLWKLGTLPPGLITFYSKTKSLDKSWHVL 503
At2g20810     GMNVFDLVEWRKRNVTGIYHYWQEKNVDR--TLWKLGTLPPGLLTFYGLTEALEASWHIL 480
At1g18580     GMNVFDLIAWRNANVTARYHYWQDQNRER--TLWKLGTLPPGLLSFYGLTEPLDRRWHVL 480
At5g54690     GMNVFDLAAWRRTNISSTYYHWLDENLKSDLSLWQLGTLPPGLIAFHGHVQTIDPFWHML 476
At3g01040     GMNIFDLRTWRKTNIRETYHSWLKENLKSNLTMWKLGTLPPALIAFKGHVQPIDSSWHML 452
At5g15470     GMNIFDLQAWRKTNIRETYHSWLRENLKSNLTMWKLGTLPPALIAFKGHVHIIDSSWHML 514
At3g58790     GMNVFDLKAWRQTNITEAYSTWLRLSVRSGLQLWQPGALPPTLAFKGLTQSLEPSWHVA 478
At1g19300     GVMVIDLSRWREGAYTSRIEEWMAMQKR--MRIYELGSLPPFLLVFAGLIKPVNHRWNQH 291
At3g50760     GVMVIELKKWREGDYTRKIIEWMELQKR--IRIYELGSLPPFLLVFAGNIAPVDHRWNQH 280
At1g13250     GVMVIDLGKWREERRVTVKLETWMRIQKR--HRIYELGSLPPFLLVFAGDVEPVEHRWNQH 284
At3g06260     GVMVVDVNKWRKGMYTQKVEEWMTIQKQ--KRIYHLGSLPPFLLIFAGDIKAVNHRWNQH 290
At1g02720     GVMVIDLKKWRQFRFTKRIEKWMEIQKI--ERIYELGSLPPFLLVFAGHVAPISHRWNQH 305
At4g02130     GVMVIDLKKWRRGGYTKRIEKWMEIQRR--ERIYELGSLPPFLLVFSGHVAPISHRWNQH 290
At3g62660     GVMVMDLERWRRVGYTEVIEKWMEIQKS--DRIYELGSLPPFLLVFAGEVAPIEHRWNQH 300
At1g24170     GVMVMDLVRWREGNYREKLEQWMQLQKK--MRIYDLGSLPPFLLVFAGNVEAIDHRWNQH 313
At1g70090     GVMVMDLVRWREGNYREKLETWMQIQKK--KRIYDLGSLPPFLLVFAGNVEAIDHRWNQH 310
At3g28340     GVMVIDLERWREGDYTRKIENWMKIQKED-KRIYELGSLPPFLLVFGGDIEAIDHQWNQH 297
At1g54940     NEIFVWWHRLSKRVNTMKYFDE--KNHRRHD-------LPENVEGLHYLG-LKPWVCYRD 478
At1g08990     NEYFVWWHRLSKRLNTMKYFGDESRHDKARN-------LPENLEGIHYLG-LKPWRCYRD 345
At4g33330     NEIFVWWHRLPRRVNFLKNFWSNTTKERNIKNNLFA-AEPPQVYAVHYLG-WKPWLCYRD 516
At1g77130     NEIFTWWHRIPKHMNFLKHFWEGDEPEIKKMKTSLFGADPPILYVLHYLGYNKPWLCFRD 502
At3g18660     NEVFTWWHRIPKHMNFLKHFWIGDEDDAKRKKTELFGAEPPVLYVLHYLG-MKPWLCYRD 534
At1g56600     NAGMFVYEPNLSTYHNLLETVKIVPPTLFAEQDFLNMYFKDIYKPIPPVYNLVLAMLWRH 244
At1g09350     NAGMFVYEPSLPTYYNLLETLKVVPPTFPAEQDFLNMYFKDIYKPIPPVYNLVLAMLWRH 238
At2g47180     NAGMFLYEPNLETYEDLLRTLKITPPTFPAEQDFLNMYFKKIYKPIPLVYNLVLAMLWRH 252
At1g60450     NAGMFVFEPNPLTYESLLQTLQVTPPTFPAEQDFLNMFFGKVFKPVSPVYNLILSVLWRH 242
At1g60470     NAGMFVFEPSPLTYESLLQTLEITPPSFPAEQDFLNMFFEKVYKPIPLVYNLVLAMLWRH 245
At4g26250     NAGMLVFEPNLLTYEDLLRVVQITTPTYFAEQDFLNEYFTDIYKPIPSTYNLVMAMLWRH 247
At5g23790     NAGMLVFGPNLVTYEDLLRVVQITTPTYFAEQDFLNIYFRDIYKPIPSTYNLVMAMLWRH 244
                .  .                                              .
```

```
At3g61130      GLG-YNPSIDKKDIEN--AAVVHYNGNMKPWLELAMSKYRPYWTKYIKFDHPYLRRCNLH 672
At2g46480      GLG-YDKGIDVKKIER--SAVIHYNGHMKPWTEMGISKYQPYWTKYTNFDHPYIFTCRLF 527
At4g38270      GLG-YDPALNQTAIEN--AAVVHYNGNYKPWLGLAFAKYKPYWSKYVEYDNPYLRRCDIN 657
At5g47780      GLG-YNPSVNQRDIER--AAVIHYNGNLKPWLEIGIPRYRGFWSKHVDYEHVYLRECNIN 614
At2g30575      GLG-HESGLRASDIEQ--AAVIHYDGIMKPWLDIGIDKYKRYWNIHVPYHHPHLQRCNIH 609
At1g06780      GLG-RESGVKAVDIEQ--AAVIHYDGVMKPWLDIGKENYKRYWNIHVPYHHTYLQQCNLQ 588
At2g38650      GLG-YDYYINAQAIKN--AAILHYNGNMKPWLELGIPNYKNYWRRHLSREDRFLSDCNVN 618
At3g25140      GLG-YNPSISMDEIRN--AAVVHFNGNMKPWLDIAMNQFRPLWTKHVDYDLEFVQACNFG 558
At3g02350      GLG-YNPGVSMDEIRN--AGVIHYNGNMKPWLDIAMNQYKSLWTKYVDNEMEFVQMCNFG 560
At2g20810      GLG-YTN-VDARVIEK--GAVLHFNGNLKPWLKIGIEKYKPLWERYVDYTSPFMQQCNFH 536
At1g18580      GLG-YDVNIDNRLIET--AAVIHYNGNMKPWLKLAIGRYKPFWLKFLNSSHPYLQDCVTA 537
At5g54690      GLGYQETTS-YADAES--AAVVHFNGRAKPWLDIAFPHLRPLWAKYLDSSDRFIKSCHIR 533
At3g01040      GLGYQSKTN-LENAKK--AAVIHYNGQSKPWLEIGFEHLRPFWTKYVNYSNDFIKNCHIL 509
At5g15470      GLGYQSKTN-IENVKK--AAVIHYNGQSKPWLEIGFEHLRPFWTKYVNYSNDFIKNCHIL 571
At3g58790      GLGSRSVKSPQEILKS--ASVLHFSGPAKPWLEISNPEVRSLWYRYVNSSDIFVRKCKIM 536
At1g19300      GLGGDNFRGLCRDLHPGPVSLLHWSGKGKPWARLDAGRPCPLDALWAPYDLLQTPFALDS 351
At3g50760      GLGGDNFRGLCRDLHPGPVSLLHWSGKGKPWVRLDDGRPCPLDALWVPYDLLESRFDLIE 340
At1g13250      GLGGDNLEGLCRNLHPGPVSLLHWSGKGKPWLRLDSRRPCPLDSLWAPYDLFRYSPLISD 344
At3g06260      GLGGDNTEGRCRTLHPGPISLLHWSGKGKPWLRLDSRKPCIVDHLWAPYDLYRSSRHSLE 350
At1g02720      GLGGDNVRGSCRDLHSGPVSLLHWSGSGKPWLRLDSKLPCPLDTLWAPYDLYKHSH---- 361
At4g02130      GLGGDNVRGSCRDLHPGPVSLLHWSGSGKPWIRLDSKRPCPLDALWTPYDLYRHSH---- 346
At3g62660      GLGGDNVRGSCRDLHPGPVSLLHWSGSGKPWFRLDSRRPCPLDTLWAPYDLYGHYSR---- 357
At1g24170      GLGGDNIRGSCRSLHPGPVSLLHWSGKGKPWVRLDEKRPCPLDHLWEPYDLYKHKIERAK 373
At1g70090      GLGGDNVRGSCRSLHKGPVSLLHWSGKGKPWVRLDEKRPCPLDHLWEPYDLYEHKIERAK 370
At3g28340      GLGGDNIVSSCRSLHPGPVSLIHWSGKGKPWVRLDDGKPCPIDYLWAPYDLHKSQRQYLQ 357
At1g54940      YDCNWDISERRVFASDSVHEKWWKVYDKMSEQLKGYCGLNKNMEKRIEKWRRIAKNNSLP 538
At1g08990      YDCNWDLKTRRVYASESVHARWWKVYDKMPKKLKGYCGLNLKMEKNVEKWRKMAKLNGFP 405
At4g33330      YDCNYDVDEQLVYASDAAHVRWWKVHDSMDDALQKFCRLTKKRRTEINWERRKARLRGST 576
At1g77130      YDCNWNVDIFQEFASDEAHKTWWRVHDAMPENLHKFCLLRSKQKAQLEWDRRQAEKGNYK 562
At3g18660      YDCNFNSDIFVEFATDIAHRKWWMVHDAMPQELHQFCYLRSKQKAQLEYDRRQAEAANYA 594
At1g56600      PENIELDQVKVVHYCAAGAKPWRFTGEEENMDREDIKMLVKKWWDIYNDESLDYKNVVIG 304
At1g09350      PENIELNEAKVVHYCAAGAKPWRFTGQEGNMEREDIKMLVEKWWDIYNDESLDYKNFNVH 298
At2g47180      PENVELGKVKVVHYCAAGSKPWRYTGKEANMEREDIKMLVKKWWDIYDDESLDYKKPVTV 312
At1g60450      PGKVDLESVKVVHYCPPGSKPWRYTGEEPNMDREDVKMLIKKWWDIYNDESLDFKPKSPA 302
At1g60470      PENVELEKVKVVHYCAAGSKPWRYTGEEANMDREDIKMLVDKWWDVYNDESLDFKSKIPA 305
At4g26250      PEHIDLDQISVIHYCANGSKPWRFDETEEHMDREDIKMLVKKWWDIYEDSSLDYKNFVET 307
At5g23790      PEHIDLDQISVVHYCANGSKPWKFDEAEEHMDREDIKMLVKKWWEIYEDSSLDYKNFVET 304
```

```
At3g61130       E---------------------------------------------------------- 673
At2g46480       E---------------------------------------------------------- 528
At4g38270       E---------------------------------------------------------- 658
At5g47780       P---------------------------------------------------------- 615
At2g30575       D---------------------------------------------------------- 610
At1g06780       A---------------------------------------------------------- 589
At2g38650       P---------------------------------------------------------- 619
At3g25140       L---------------------------------------------------------- 559
At3g02350       L---------------------------------------------------------- 561
At2g20810       ----------------------------------------------------------
At1g18580       ----------------------------------------------------------
At5g54690       AS--------------------------------------------------------- 535
At3g01040       E---------------------------------------------------------- 510
At5g15470       E---------------------------------------------------------- 572
At3g58790       N---------------------------------------------------------- 537
At1g19300       ----------------------------------------------------------
At3g50760       S---------------------------------------------------------- 341
At1g13250       S---------------------------------------------------------- 345
At3g06260       E---------------------------------------------------------- 351
At1g02720       ----------------------------------------------------------
At4g02130       ----------------------------------------------------------
At3g62660       ----------------------------------------------------------
At1g24170       DQSLLGFASLSELTDDSSFL---------------------------------------- 393
At1g70090       DQSLFGFSSLSELTEDSSFF---------------------------------------- 390
At3g28340       YNQELEIL---------------------------------------------------- 365
At1g54940       DRHWEIEVRDPRKTNLLVQ----------------------------------------- 557
At1g08990       ENHWKIRIKDPRKKNRLSQ----------------------------------------- 424
At4g33330       DYHWKINVTDPRRRRSYLIG---------------------------------------- 596
At1g77130       DGHWKIKIKDKRLKTCFEDFCFWESMLWHWGETNSTNNSSTTTTSSPPHKTALPSL---- 618
At3g18660       DGHWKIRVKDPRFKICIDKLCNWKSMLRHWGESNWTDYESFVPTPPAITVDRRSSLPGHN 654
At1g56600       -----DSHKKQQTLQQFIEALSEAGALQYVKAPSAA----------------------- 335
At1g09350       CGQKEDVHRKPKTLPQFFTDLSEADVLQCAKAPSAA----------------------- 334
At2g47180       V----DTEVDLVNLKPFITALTEAGRLNYVTAPSAA----------------------- 344
At1g60450       D-------LEATVLESTIIASVTEAPLSYSPAAPSAA----------------------- 332
At1g60470       D-------AEETVTKSSILASVLEPEMTYFPAPSAA----------------------- 334
At4g26250       ES-------KLSPINATLASKESVGDVLISLAPSAA----------------------- 336
At5g23790       ES-------KLNPVTATLASKKLVGDVLTSLAPSAA----------------------- 333
```

At3g61130	-
At2g46480	-
At4g38270	-
At5g47780	-
At2g30575	-
At1g06780	-
At2g38650	-
At3g25140	-
At3g02350	-
At2g20810	-
At1g18580	-
At5g54690	-
At3g01040	-
At5g15470	-
At3g58790	-
At1g19300	-
At3g50760	-
At1g13250	-
At3g06260	-
At1g02720	-
At4g02130	-
At3g62660	-
At1g24170	-
At1g70090	-
At3g28340	-
At1g54940	-
At1g08990	-
At4g33330	-
At1g77130	-
At3g18660	L 655
At1g56600	-
At1g09350	-
At2g47180	-
At1g60450	-
At1g60470	-
At4g26250	-
At5g23790	-